AD__________________

Award Number: DAMD17-00-1-0269

TITLE: Biological Function of Matriptase, a Novel Trypsin-like Protease, in Human Breast Cancer

PRINCIPAL INVESTIGATOR: Sheau-Ling Lee, Ph.D.

CONTRACTING ORGANIZATION: Georgetown University Medical Center
Washington, DC 20057

REPORT DATE: July 2001

TYPE OF REPORT: Annual Summary

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 2001	Annual Summary (1 Jul 00 - 30 Jun 01)

4. TITLE AND SUBTITLE
Biological Function of Matriptase, a Novel Trypsin-like Protease, in Human Breast Cancer

5. FUNDING NUMBERS
DAMD17-00-1-0269

6. AUTHOR(S)

Sheau-Ling Lee, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Georgetown University Medical Center
Washington, DC 20057

email - lees13@georgetown.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
Report contains color

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. Abstract ***(Maximum 200 Words)*** *(abstract should contain no proprietary or confidential information)*

Matriptase is an epithelial-derived, type 2, integral membrane, serine protease. It contains an N-terminal putative transmembrane sequence, followed by multiple LDL and CUB repeats, and the C-terminal protease domain. To further understand its role in breast cancer, I have investigated its substrate specificity and its interaction with inhibitors. I showed in this report that matriptase specifically cleaves synthetic peptide after an Arg or Lys residue, and prefers small side-chain amino acid, such as Ala and Gly, at P2 site. A 3-D structure of the protease domain of matriptase was built by the homology modeling. This modeled structure was used in a structure-based screening of inhibitors. Two bis-benzamindine derivatives were developed to be both potent and selective inhibitor of matriptase. In addition, a naturally occurring trypsin inhibitor from sunflower seed was found to inhibit matriptase as well as HAI1; this inhibitor had little inhibition to thrombin and uPA. I have also shown that matriptase can activate HGF and pro-uPA. These results further support the hypothesis that matriptase acts as an upstream activator in metastasis and cancer invasion by interacting with and recruiting various factors to the site of contact between cancer and stromal cells; and by degrading or processing a broad range of substrates.

14. SUBJECT TERMS
breast cancer, protease. trypsin

15. NUMBER OF PAGES
44

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

INTRODUCTION: Extracellular proteolytic systems have been proposed to play major roles in cancer cell invasion and metastasis [13]. They degrade the extracellular matrix, the basement membrane, the basal laminae, and the interstitial stoma allowing cancer cells migrate to adjacent cells [13]. Currently, two proteolytic systems, the urokinase plasminogen activator (uPA)/uPA-receptor /plasminogen and the zinc-dependent metalloproteinases, have been proposed to be responsible for the majority of proteolysis of pericellular [13]. However, both systems are synthesized primarily in the tumor stroma, and require indirect mechanisms for their recruitment and activation on the surfaces of carcinoma cells [13]. We have recently characterized a novel, integral membrane serine protease (matriptase) and Kunitz-type inhibitor system in human breast cancer [9-11]. The mouse homologue of matriptase was also cloned and termed epithin [7]. In contrast to the urokinase plsasminogen activator (uPA)/uPA-receptor /plasminogen and the zinc-dependent metalloproteinases, both matriptase and its endogenous Kunitz-type inhibitor are expressed by cultured breast epithelial cells and cancer cells. The open reading frame of matriptase/epithin contains a trypsin-like serine protease domain on the C-terminal half. The N-terminal non-catalytic region of matriptase/epithin contains two tandem repeats of complement subcomplements 1r, 1s (C1r/s) and four tandem repeats of LDL receptor domains that are likely to be involved in protein-protein interaction. In addition, a 14-amino acid hydrophobic region that is likely to serve as signal peptide and a putative transmembrane domain [26] were found at the N-terminus. Kunitz-type inhibitor of matiptase was initially identified as an inhibitor of hepatocyte growth factor activator (HAI1). The cDNA sequences of HAI1 contains a transmembrane domain at its C-terminus and two kunitz domains separated by a LDL receptor domain. These domains are clustered in the C-terminal half of the protein. I proposed that matriptase/epithin induces metastasis and cancer invasion by interacting with and recruiting various factors to the site of contact between cancer and stromal cells; and by degrading or processing a broad range of substrates. This project is designed to examine in detail the biochemical characteristics, the biological substrates, and the regulation of matriptase/epithin activity.

PROGRESS MADE IN THIS REPORT: *Aim 1. Determine the steady-state kinetic parameters for matriptase and the inhibition constant for HAI1 and other serine protein inhibitors.* **I have completed most of the works in this aim. However, a small part of the studies is still under investigatoin.** *Aim 2 Identify potential biological substrates for matriptase*. **I have completed this aim.** *Aim 3. Examine the mechanisms for matriptase activation and the role of the non-catalytic region in this processes.* **I request a revision in this aim to now investigate the involvement of matriptase/epithin in the normal mammary gland development and in the breast cancer progression.**

BODY: In the period of July 2000-June 2001, I have addressed most of *Aim 1. Determine the steady-state kinetic parameters for matriptase and the inhibition constant for HAI1 and other serine protein inhibitors* and the entire *Aim 2 Identify potential biological substrates for matriptase*. Studies in aim 2 and the part of kinetic studies of matriptase in aim 1 have been published in the **Journal of Biological Chemistry in an article titled "Activation of hepatocyte growth factor and urokinase/plasminogen acitvator by matriptase, an epithelial membrane serine protease"** (reprint included). In these studies, the active matriptase enzyme was purified form human milk using a combination of CM-sepharose and immunoaffinity. By measuring the Km and Vmax using various fluorescent substrate peptides, the substrate specificity of matriptse was determined. As shown in table 1 in the paper, matriptae selectively cleaves peptide after an Arg or Lys residue. Matriptase prefers to bind to peptides containing small side-chain amino acids, such as Ala and Gly, at P2 site; and the binding affinity to the former is about 30-fold higher than that to the later. The most reactive substrate peptides for matriptase are N-ter-butoxycarbonyl-Gln-Ala-Arg-7-amino-4-methyl-coumarin, with a Km of 4.98 μM, and N-ter-butoxycarbonyl-γ-benzyl-

Glu-Ala-Arg-7-amino-4-methyl-coumarin, with a Km of 3.81 μM. Matriptase cleaves the former peptide substrate with Vmax about 7-fold higer the later peptide substrate. These two peptide substrates were reported to be good substrates for bovine trypsin and human factor XIa, respectively. In figures 1 to 3 in the paper, I showed that matriptase can proteolytically cleave the latent form, single-chain hepartocyte growth factor (HGF) into a two-chain form that contains the α- and β-chain of HGF. The matriptase-cleaved HGF can stimulate scattering of the Madin-Darby canine kidney epithelial cell line and tyrosine phosphorylation of the receptor c-Met in the A549 human lung carcinoma cell line. These data evidence that matriptase cleavage of HGF is a process of HGF activation. Also shown in this paper is the observation that matripatase can proteolytically convert pro-uPA into the protease active, two-chain form of uPA (figure 5). Interestingly, matriptase does not cleave plasminogen, a protein that shears high homology with HGF and also requires cleavage at Arg for activation. Under the same assay condition, matriptase cleavage of extracellular matrix protein, such as laminin, fibronectin, collogenes was not observed either (Lee, unpublished observations).

For the second part of the studies in *Aim1.* _the inhibition constant for HAI1 and other serine protein inhibitors_, we have finished two manuscripts. In these studies, the fluorescent peptide N-ter-butoxycarbonyl-Gln-Ala-Arg-7-amino-4-methyl-coumarin was used as a matriptase substrate. HAI1 protein used in these studies was purified from T47D human breast cancer cell. These preliminary studies showed that the inhibition constant (Ki) of HAI1 to matriptase is in the 1 nM range. For studies of other inhibitors, I started with two approaches, one is to examine the natural occurring inhibitor peptides, and another is to search small compound inhibitors. In our publication **"Structure-based approach for the discovery of bis-benzamidines as novel inhibitors of matriptase"** in the **Journal of Medicinal Chemistry** (Reprint included.), we described the screening of small compound inhibitors of matriptase using molecular modeling strategy. In this approach, the 3-D structure of the protease domain of matriptase was built using the x-ray structure of thrombin as template. Since there is a high amino acid homology between matriptase and thrombin, the modeled structure of matriptase should be very accurate. This structure was used in structure-based screening in the NCI small compound database. From the 2000 best scoring compounds one group of inhibitors identified is bis-benzamindines. From the seven bis-benzamindine analogs available from NCI database, we have identified two derivatives to be both potent and selective: they inhibit matriptase with Ki around 200 nM and they inhibit uPA and thrombin less efficiently. The first natural inhibitor peptide that attract my attention is Bowman-Birk inhibitor (BBI), a major inhibitor in soybean. BBI has been demonstrated to be anti-carcinogenic both in vivo and in vitro. However, a target protease of BBI has not been identified. It is very interesting to examine if matriptase is a possible membrane protease for BBI. Therefore, the ability of BBI inhibit matriptase was examined. The molecular modeling approach used in the small compound screening also revealed another possible natural peptide inhibitor for matriptase. This is a recently discovered trypsin inhibitor from sunflower seeds, termed sunflower trypsin inhibitor (SFTI). The natural product of SFTI, purified from sunflower seeds, is a 14-aa backbone-cyclized peptide that is further stabilized by an intramolecular cystine disulfide bond [12]. The pilot modeled structure of SFTI-matriptase complex (not shown, but included in figure 1 in the included manuscript) suggested that the inhibitory loop of SFTI binds to matriptase as well as, if not better than, HAI1. As shown in our manuscript in press in **Bioorganic and Medicinal Journal titled "Synthesis and evaluation of the sunflower derived trypsin inhibitor as a potent inhibitor of the type II transmembrane serine protease, matriptase"** (manuscript included), SFTI inhibited matriptase enzyme activity with a Ki of 0.92 nM. In addition to its potency, SFTI appears to be highly selective. It has little inhibition for thrombin, and no inhibition for uPA (table 1 in the manuscript). BBI also appeared to be a good inhibitor for matriptase, although less efficient than SFTI (figure 1 in the manuscript).

Part of the inhibitor studies for aim 1 are now ongoing. The main focus in these studies is to compare the action between the Kunitz domain I, II of HAI1, the trypsin-inhibitor loop of BBI, and inhibitory loop of SFTI. I have set up the baculovirus/insect cell expression system for HAI1, the protease domain and the full-length of matriptase (figure). Various Kunitz domain mutants of HAI1 have been made by *in vitro* mutagenesis, and are ready to be put into baculovirus/insect expression system. Synthetic peptides containing the trypsin inhibitory loop of BBI, the Kunitz I loop, and the Kunitz II loop of HAI1 have been made. SFTI analogs with possible improved stability and potency also have been synthesized. I am now working out the preparation of fully active protease domain of matriptase. These studies will provide us with more detailed understanding about the inhibitor-matriptase/epithin interaction. Some of these inhibitors may become useful tools for studies of the function of matriptase/epithin *in vivo* as well as potential therapeutic means.

Aim 3. Examine the mechanisms for matriptase activation and the role of the non-catalytic region in this processes. Based on the following reasons, I am requesting a revision of this specific aim to **Investigate the involvement of matriptase/epithin in the normal mammary gland development and in the breast cancer progression**. As presented in this report, matriptase can act as activator of HGF and pro-uPA. uPA is an important activator in extracellular matrix degradation. HGF has been shown to have multifunctional effects on mammalian cells. These include mitogenensis [4;5;15;17;24;30], motogenesis [4;23], mophorgenesis [6;14;19], and recently tumor angiogenesis [20;21]. HGF/c-Met signalling has been found to be required for appropriate mammary gland development [18], and for normal embroyonic development [3;22;28] in mice. De-regulated HGF/c-Met signalling is also associated with a variety of human cancers, including breast cancer [1;8;16;27]. HGF transgenic mice exhibit tumorigenesis in a wide variety of tissue types including mammary gland [25]. Thus, matriptase and its mouse homolog epithin provide a perfect link in cell migration that they stimulate the transition of epithelial cells to mobile, fibroblast-like cells, and activate extracelluar matrix degradation to open the pathways for the migrating cells.

Recently, it was found that in the non-tumorigenic human mammary epithelial 184A1N4 and MCF-10A cell lines, the activated form of matriptase can be stimulated by phospholipid lysophosphatidic acid and sphingosine-1-phosphate [2] (and Benaud et al. manuscript is in preparation). While in human breast cancer cell lines including T47D, MCF-7, and MDA-MB-468, matriptase is constitutively in the activated form and the activation is not regulated by phospholipid lysophosphatidic acid or sphingosine 1 phosphate [2] (and Benaud et.al, manuscript in preparation). It is not clear how these lipids involve in the activation or what the real mechanism is in the activation. These observations, however, suggested that the activity of matriptase needs to be tightly controlled in a normal condition, and that cancer cells might have developed a system that allow them to escape from the normal regulation of matriptase activity. How this system, assuming this is the case in cancer cells, is developed is another mystery.

Obviously, to further understand the role of matriptase/epithin in breast cancer progression, it is needed to first understand what its function is in the normal mammary. Further understanding of its changes and/or the induced changes in the tumor will then can be used to develop strategies in cancer occurrence prevention. Taking into account the fact that matriptase can activate HGF and pro-uPA, **I hypothesize that matriptase/epithin play a role in the normal mammary gland development and it could direct the development of breast tumor in an unfavorable physiological condition.** I propose to re-focus my specific aim 3 from "*Examine the mechanisms for matriptase activation and the role of the non-catalytic region in this processes*" to "*Investigate the involvement of matriptase/epithin in the normal mammary gland development and in the breast cancer progression*". Considering the availability and the power of normal and transgenic rodent models, I will extend these studies to the mice systems. *Aim 3a, I will examine the significance of epithin/matritpase activity on the morphorgeneis of mammary epithelium.* It has been

shown in the mouse mammary epithelial cell line EpH4/K6, and in both mouse and human primary mammary epithelial cultures, that they can be stimulated by HGF to grow into ductal structures in matrigel [18;19;29]. I have received the EpH4/K6 cells from Dr. Birchmeier (Max-Planck Institute, Berlin), and the anti-epithin antibodies from Dr. Park (Seoul National Univ., Seoul). Using these antibodies, I have detected the expression of epithin in EpH4/K6 cells and in the mammary (figures 1 and 2). The two anti-epithin antibodies used in these pilot experiments appears to recognize either the full-length and the N-terminal fragment of the cleavage-activated epithin (figure 1, "p"), or the protease domain of the cleavage-activated epithin (figure 1, "m"). EpH4/K6 cells expressed mostly the full-length, uncleaved epithin (figure 1, A), the mostly cleaved, two-chain epithin (figure 1, C), and both forms of epithin (figure 1, B) were observed. Apparently, the active state of epithin in EpH4/K6 cell seems also to be regulated under culture condition. Thus, this system appears to be a good, simple cell system to study the involvement of epithin in the mammary gland growth. The matriptase inhibitors developed in this report will be used to in the studies to inhibit epithin protease activity. I will also do the studies in primary mammary cultures since these cultures can be used later in animal systems. In addition to the normal primary culture, I will also make primary culture from HGF mammary. It was reported that this tissue expresses a high level of HGF, it is prone to tumor formation, and it exhibits developmental abnormalities [25]. The HGF mammary tumor also expresses epithin (figure 2). Studies in this system could be compared with the studies in normal mammary system, providing further clues to how epithin/matriptase involves in mammary development. *Aim 3b, I will examine the expression pattern and the localization of matriptase/ epithin during mammary gland development in the normal mice and in the HGF transgenic mice*. For these studies, the mammary gland will be taken from mice at deferent ages, the expression of epithin will be examined with Western immunoblot, and the localization of epithin will be examined with immunohistochemical staining of the tissue sections. The same studies will be repeated in HGF transgenic mice. Collections from these studies might shed some light on the mechanism of matriptase/epithin in tumorigenesis. Currently, there is no antibody that can be used to investigate activated form of epithin in tissue sections. However, with the two anti-epithin antibodies mentioned above, it may be possible to examine the extent of activated epithin by Western immunoblot of proteins prepared in the absence or presence of reducing agent. In *Aim 3c, I will examine the influence of the matriptase inhibitors developed in aim 2 in these processes*.

KEY RESEARCH ACCOMPLISHMENTS

In the first year, I have:

1. Shown that, using peptide substrates in enzymatic assay, matriptase exhibits trypsin-like activity, selectively cleaves synthetic peptides after an Arg or Lys residue. In addition, matriptase prefers peptides containing small side-chain amino acids, such as Ala and Gly, at P2 site. The most reactive substrate peptide for matriptase is N-ter-butoxycarbonyl-Gln-Ala-Arg-7-amino-4-methyl-coumarin, with a Km of 4.98 μM and the highest Vmax.
2. Shown that matriptase can convert HGF and pro-uPA proteolytically into their functional, active, two-chain form.
3. Shown that matriptase does not cleave plasminogen or other extracellular matrix protein.
4. Built a 3-D structure of the protease domain of matriptase, and used this modeled structure successfully in the inhibitor screening.
5. Developed dibenzamidine derivatives as potent small compound inhibitor for matriptase, and found two analogs that are both potent and selective.

6. Shown that a small, cyclic peptide trypsin inhibitor from sunflower seed, SFTI, can inhibit matritpase activity as well as the HAI1 does. I have also shown that this peptide inhibitor has little or no inhibition to thrombin and uPA.

REPORTABLE OUTCOMES

Publications

1. Lee, S.-L., Dickson, R.B., and Lin, C.-Y., 2000, Activation of hepatocyte growth factor and urokinase type plasminogen activator by matriptase, an epithelial membrane serime protease. J. Biol. Chem. 275, 36720-36725.

2. Enyedy, I.J., Lee, S.-L., Kuo, A.H., Dickson, R.B., Lin, C.-Y., and Wang, S., 2001, Structure-based approach for the discovery of bis-benzamidines as novel inhibitors of matriptase. J. Medicinal Chem. 44, 1349-1355.

3. Long, Y.-Q., Lee, S.-L., Lin, C.-Y., Enyedy, I., Wang, S., Dickson, R.B., and Roller P.P., 2001 Synthesis and evaluation of the sunflower derived trypsin inhibitor as a potent inhibitor of the type II transmembrane serine protease, matriptase. Bioorg. Med. Chem. Lett. In press.

4. Roller, P.P., Long, Y.-Q., Li, P.; Lee, S.-L., Lin, C.-Y., Enyedy, I.;Wang, S., Dickson, R.B., 2001, Bicyclic peptide inhibitors of an epithelial cell-derived transmembrane protease, matriptase. In 'Peptides: The Wave of the Future' (Proceedings of the 2nd International / 17th American Peptide Symposium), R.A. Houghten and M. Lebl (Eds.). In press.

Abstracts

Lee, S.-L., Dickson, R. B., and Lin, C.-Y., 2000, Biological function of matriptase, a novel trypsin-like protease. FACS/The Salk Institute for biological studies. 16th Annual Meeting on Oncogenes and tumor suppressor. La Jolla, CA

Enyedy, I., Lee, S.-L., Lin, C.-Y., Dickson, R. B., and Wang, S., 2000, Structure based design of inhibitors for matriptase. Proceeding of the Annual meeting of the American Chemical Society. Washington, DC

Roller, P. P., Long, Y.-Q., Li, P., Lee, S.-L., Lin, C.-Y., Enyedy, I., Wang, S., and Dickson, R. B., 2001, Bicyclic peptide inhibitors of an epithelial cell-derived transmembrane protease, matriptase. Proceedings of the 17th American Peptide Symposium. San Diego, CA

CONCLUSIONS: In the first year, **I have completed most of the studies in *Aim 1:. Determine the steady-state kinetic parameters for matriptase and the inhibition constant for HAI1 and other serine protein inhibitors* and *Aim 2: Identify potential biological substrates for matriptase*.** Using a molecular modeling strategy in a structure-based screening, we have developed bis-benzamidine analogs the potent and selective small compound inhibitors for matriptase. We have found a naturally occurring trypsin inhibitor from sunflower seed, SFTI, to be the most potent and selective matriptase inhibitor tested. These inhibitors are suitable for used in cell culture since they do not posses significant toxicity to the cultured cell lines (figure 3). These inhibitors have low inhibitory activities against uPA and thrombin, the serine proteases involving in the exptracellular matrix degradation and in the blood coagulation, respectively. The constrained, conformational rigid structure of these inhibitors further provides a promising lead template for further pharmacophore-based

inhibitor development. In addition, the success in our inhibitor screening proves that molecular modeling is a useful approach in searching for an inhibitor of a protein without an available X-ray structure. **I will continue inhibitor studies to compare the actions of the Kunitz domain I, II of HAI1, the trypsin-inhibitor loop of BBI, and the inhibitory loop of SFTI.** These studies will provide us with a detailed understanding about inhibitor-matriptase/epithin interactions. These inhibitors may become useful tools for studies of the function of matriptase/epithin *in vivo*. In addition, they may provide a new therapeutic approach.

I have shown in this report that matriptase can proteolytically convert HGF and pro-uPA into their active forms. Matriptase, however, does not cleave plasminogen or other extracellular matrix proteins that we have studied to date. These results further support the hypothesis that matriptase/epithin could function as an upstream, epithelial membrane activator to recruit and activate stromal-derived downstream effectors important for extracellular matrix degradation and epithelial migration, two major events of tissue remodeling, cancer invasion, and metastasis. To further understand the biological function of matriptase/epithin, and the possible mechanism(s) of its involvement in breast cancer progression, **I will extend my studies to the mouse system to investigate the role of epithin in normal mammary gland development and its expression in the mammary tumor.**

Figure 1. Expression of epithin in EpH4/K6 cells. EpH4/K6 cells were grown in DME medium supplemented with 10% fetal calf serum until confluence. Cells were either extracted immediately (B), or transferred to fresh medium at low density (A) or high density (C), and then extracted when they were confluent (1 day in C, and 4 days in A). Equal amounts of protein were boiled in Laemmli sample buffer, separated by SDS-PAGE, transferred to membrane, and probed with the polyclonal (p) or the monoclonal (m) anti-epithin antibody. The full-length epithin is about 95 Kdal detected by the polyclonal antibody. Activation cleavage of eiptihin resulted a protease domain of about 31 Kdal (arrow head) and a fragment runs about 66 Kdal that were detected by the monoclonal and the polyclonal anti-epithin, respectively.

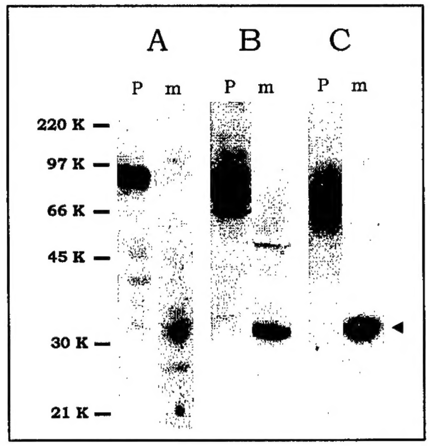


Figure 2. Expression of epithin in the mammary tissues. Sample of proteins extracted from the normal mouse mammary gland (a), the mammary tumor from c-myc transgenic mouse (b), or HGF transgenic mouse (c), were treated with sample buffer without (reducing reagent, and immunoblotted with the polyclonal anti-epithin antibody as described in figure 1.

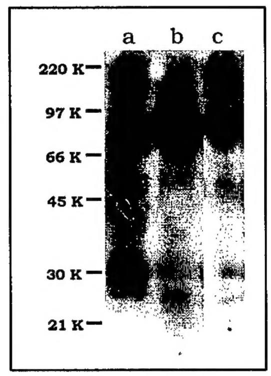
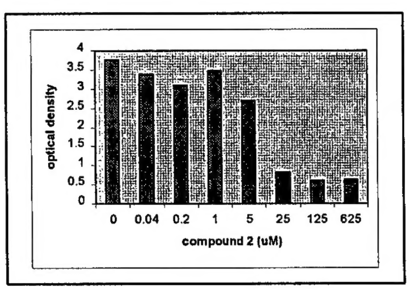
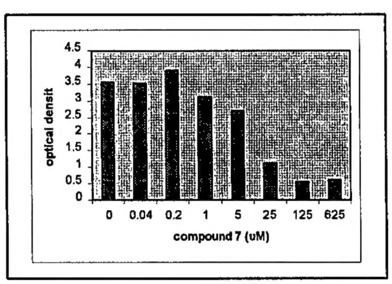
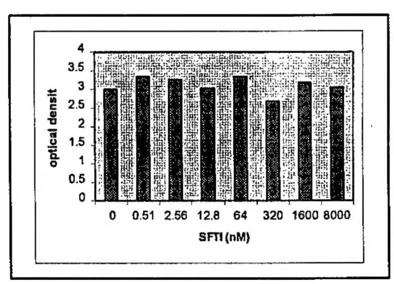


Figure 3. Toxicity of bis-benzamidine compound inhibitors and SFTI in cell culture.

To each well in the 96-well culture plate was added 100 ul DME medium supplemented with 10% fetal calf serum in the absence or the presence of different concentration of the compound inhibitors or SFTI. About 3000 EpH4/K6 cells in 150 ul of medium were added to each well, and the plate was placed in a 37 C incubator until the cells in the control reached confluence. Cells were then stained with crystal violet, the crystal violet on the cells were then dissolved with sodium citrate and read with a plate reader at 570 nm. Noted that the Ki values for compound 1, 7, and SFTI are 191 nM, 208 nM, and 0.92 nM, repectively.







REFERENCE

1. Altstock, R. T., Stein, G. Y., resau, J. H., and Tsarfaty, I. (2000) *Cytometry* **41,** 155-165

2. Benaud, C., Dickson, R. B., and Lin, C. Y. (2001) *European Journal of Biochemistry* **268,** 1439-1447

3. Bladt, F., Riethmacher, D., Isenmann, S., Aguzzi, A., and Birchmeier, C. (1995) *Nature* **376,** 768-771

4. Gherardi, E. and Stoker, M. (1991) *Cancer Cells* **3,** 227-232

5. Gohda, E., Tsubouchi, H., Nakayama, H., Hirono, S., Sakiyama, O., Takahashi, K., Miyazaki, H., Hashimoto, S., and and Daikuhara, Y. (1988) *J.Clin.Invest.* **81,** 414-419

6. Kamalati, T., Niranjan, B., Yant, J., and and Buluwela, L. (1999) *J.Mamm.Gland Biol.Neoplasia* **4,** 69-77

7. Kim, M. G., Chen, C., Lyu, M. S., Cho, E. G., Park, D., Kozak, C., and Schwartz, R. H. (1999) *Immunogenetics* **49,** 420-428

8. Lee, J. H., Han, S. U., Cho, H., Jennings, B., Gerrard, B., Dean, M., Schmidt, L., Zbar, B., and Vande Woude, G. F. (2000) *Oncogene* **19,** 4947-4953

9. Lin, C.-Y., Anders, J., Johnson, M., and Dickson, R. B. (1999) *J.Biol.Chem.* **274,** 18237-18242

10. Lin, C.-Y., Anders, J., Johnson, M., Sang, Q. A., and Dickson, R. B. (1999) *J.Biol.Chem.* **274,** 18231-18236

11. Lin, C.-Y., Wang, J.-K., Torri, J., Dou, L., Sang, Q. A., and Dickson, R. B. (1997) *J.Biol.Chem.* **272,** 9147-9152

12. Luckett, S., Garcia, R. S., Barker, j. j., Konarev, A. V., Shewry, P. R., Clarke, A. R., and Brady, R. L. (1999) *Journal of Molecular Biology* **290,** 525-533

13. Mignatti, P. and. Rifkin, D. B. (1993) *Physiol.Rev.* **73,** 161-195

14. Montesano, R., Schaller, G., and Orci, L. (1991) *Cell* **66,** 697-791

15. Nakamura, T., Nawa, K., Ichihara, A., Kaise, N., and and Nishino, T. (1987) *FEBS Lett.* **224,** 311-316

16. Nakopoulou, L., Gakiopoulou, H., Keramopoulos, A., Giannopoulou, I., Athanassiadou, P., Mavrommatis, J., and Davaris, P. S. (2000) *Histopathology* **36,** 313-325

17. Naldini, L., Tamagnone, L., Vigna, E., Sachs, M., Hartmann, G., Birchmeier, W., Daikuhara, Y., Tsubouchi, H., Blasi, F., and and Comoglio, P. M. (1992) *EMBO J.* **11,** 4825-4833

18. Niemann, C., Brinkmann, V., Spitzer, E., Hartmann, G., Sachs, M., Naundorf, H., and Birchmeier, W. (1998) *Journal of Cell Biology* **143,** 533-545

19. Niranjan, B., Buluwela, L., Yant, J., Perusinghe, N., Atherton, A., Phippard, D., Dale, T., Gusterson, B., and and Kamalati, T. (1995) *Development* **121,** 2897-2908

20. Rosen, E. M., Lamszus, K., Laterra, J., Polverini, P. J., Rubin, J. S., and Goldberg, I. D. (1997) *Ciba Foundation Symposium* **212,** 215-226

21. Rosen, E. M., Zitnik, R. J., Elias, J. A., Bhargava, M. M., Wines, J., and Goldberg, I. D. (1993) *EXS* **65,** 301-310

22. Schmidt, C., Bladt, F., Goedecke, S., Brinkmann, V., Zschiesche, W., Sharpe, M., Gherardi, E., and Birchmeier, C. (1995) *Nature* **373,** 699-702

23. Stoker, M. G. E., Perryman, M., and and Gray, J. (1987) *Nature* **327,** 239-242

24. Strain, A., Ismail, T., Tsubouchi, H., Arakaki, N., Hishida, T., Kitamura, N., Daikuhara, Y., and and McMaster, P. (1991) *J.Clin.Invest.* **87,** 1853-1857

25. Takayama, H., La Rochelle, W. J., Otsuka, T., anver, M., Aaronson, S. A., and Merlino, G. (1997) *Proc.Natl.Acad.Sci.U.S.A.* **92,** 701-706

26. Takeuchi, T., Shuman, M. A., and Craik, C. S. (1999) *Proceedings of the National Academy of Sciences of the United States of America* **96,** 11054-11061

27. Tsarfaty, I., Alvord, W. G., resau, J. H., Altstock, R. T., Lidereau, R., Bieche, I., Bertrand, F., Horev, J., Klabansky, R. L., Keydar, I., and Vande Woude, G. F. (1999) *Analytical & Quantitative Cytology & Histology* **21,** 397-408

28. Uehara, Y., Minowa, O., Mori, C., Shiota, K., Kuno, J., Noda, T., and Kitamura, N. (1995) *Nature* **373,** 702-705

29. Yant, J., Buluwela, L., Niranjan, B., Gusterson, B., and Kamalati, T. (1998) *Experimental Cell Research* **241,** 476-481

30. Zamegar, R. and and Michalopoulos, G. (1989) *Cancer Res.* **49,** 3314-3320

APPENDICES

Reprints and Manuscript:

1. Lee, S.-L., Dickson, R.B., and Lin, C.-Y., 2000, Activation of hepatocyte growth factor and urokinase type plasminogen activator by matriptase, an epithelial membrane serime protease. J. Biol. Chem. 275, 36720-36725.

2. Enyedy, I.J., Lee, S.-L., Kuo, A.H., Dickson, R.B., Lin, C.-Y., and Wang, S., 2001, Structure-based approach for the discovery of bis-benzamidines as novel inhibitors of matriptase. J. Medicinal Chem. 44, 1349-1355.

3. Long, Y.-Q., Lee, S.-L., Lin, C.-Y., Enyedy, I., Wang, S., Dickson, R.B., and Roller P.P., 2001, Synthesis and evaluation of the sunflower derived trypsin inhibitor as a potent inhibitor of the type II transmembrane serine protease, matriptase. Bioorg. Med. Chem. Lett. In press.

THE JOURNAL OF BIOLOGICAL CHEMISTRY

Vol. 275, No. 47, Issue of November 24, pp. 36720–36725, 2000
Printed in U.S.A.

Activation of Hepatocyte Growth Factor and Urokinase/Plasminogen Activator by Matriptase, an Epithelial Membrane Serine Protease*

Received for publication, August 25, 2000
Published, JBC Papers in Press, August 28, 2000, DOI 10.1074/jbc.M007802200

Sheau-Ling Lee‡, Robert B. Dickson, and Chen-Yong Lin§

From the Department of Oncology, Lombardi Cancer Center, Georgetown University, Medical Center, Washington, DC 20007

Matriptase is an epithelial-derived, integral membrane serine protease. The enzyme was initially isolated from human breast cancer cells and has been implicated in breast cancer invasion and metastasis. In the current study, using active matriptase isolated from human milk, we demonstrate that matriptase is able to cleave various synthetic substrates with arginine or lysine as their P1 sites and prefers small side chain amino acids, such as Ala and Gly, at P2 sites. For the most reactive substrates, *N-tert*-butoxycarbonyl (*N-t*-Boc)-γ-benzyl-Glu-Ala-Arg-7-amino-4-methylcoumarin (AMC) and *N-t*-Boc-Gln-Ala-Arg-AMC, the K_m values were determined to be 3.81 and 4.89 μM, respectively. We further demonstrated that matriptase can convert hepatocyte growth factor/scattering factor to its active form, which can induce scatter of Madin-Darby canine kidney epithelial cells and can activate c-Met tyrosine phosphorylation in A549 human lung carcinoma cells. In addition, we noted that matriptase can activate urokinase plasminogen activator but has no affect on plasminogen. These results suggest that matriptase could act as an epithelial, upstream membrane activator to recruit and activate stromal-derived downstream effectors important for extracellular matrix degradation and epithelial migration, two major events of tissue remodeling, cancer invasion, and metastasis.

Tissue remodeling is observed both in physiological and pathologic processes. These include organ development, morphogenesis, wound healing, cancer invasion, and metastasis. Degradation of extracellular matrix (ECM)[1] and cellular migration are two prominent steps in tissue remodeling. Considering that the majority of the ECM-degrading proteases, such as the plasmin/urokinase type plasminogen activator system (1), and the major motility factor, hepatocyte growth factor (HGF)/scattering factor (SF) (2) are mainly produced by the stromal components *in vivo*, tissue remodeling is likely to be an event that depends entirely upon stromal-epithelial collaboration (3). A search for epithelial-derived proteases, which may interact both with stromal-derived ECM-degrading protease systems and with motility factors, could provide a missing link in our understanding of tissue remodeling and cancer invasion and metastasis.

To investigate the epithelial role in tissue remodeling and in many aspects of tumor behavior, including growth and metastasis, we have carried out our studies on an epithelial-derived, integral membrane, trypsin-like, serine protease (matriptase) and its cognate, Kunitz-type serine protease inhibitor (hepatocyte growth factor activator inhibitor-1, HAI-1) (4–7). Matriptase is a type 2, integral membrane, trypsin-like serine protease with two putative regulatory modules: two tandem repeats of a CUB (C1r/s, Uegf, and Bone morphogenetic protein-1) domain and four tandem repeats of a low density lipoprotein (LDL) receptor domain (also see updated sequence in the GenBank™/EBI Data Bank with accession number AF118224). Matriptase was independently cloned by others, and termed membrane-type serine protease 1 (MT-SP1) (8). The mouse homologue of matriptase was also cloned and termed epithin (9). The cognate inhibitor of matriptase is a type 1 integral membrane protein, containing two Kunitz domains, separated by an LDL receptor domain (7). The inhibitor was independently characterized by others as an inhibitor (HAI-1) of hepatocyte growth factor activator, an enzyme identified in serum (10).

Considering that matriptase exhibits trypsin-like activity and presents on the surfaces of epithelial cells, and that activation of the uPA system and HGF/SF requires cleavage at Arg or Lys, we hypothesize that matriptase could act as an upstream, epithelial membrane activator of the downstream, stromal-derived effectors of tissue remodeling. In the current study, we set out to investigate the potential collaboration between epithelial and stromal cells by examining if matriptase is able to activate HGF/SF and the protease components of the uPA system. Using the 70-kDa active matriptase, purified from human milk, we are able to demonstrate that matriptase can activate pro-uPA and pro-HGF, but not plasminogen. These results reveal that a novel mechanism involving both an upstream epithelial membrane activator as well as downstream stromal effectors may play an important role in tissue remodeling.

MATERIALS AND METHODS

Antibodies—Polyclonal antibodies to hepatocyte growth factor (HGF) α-chain (C-20) and β-chain (N-19) were purchased from Santa Cruz

* Supported in part by National Institutes of Health, Specialized Program of Research Excellence Grant 1P50CA58158 in breast cancer and National Institutes of Health Grant R21-CA80897. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *"advertisement"* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Supported by Department of Defense Fellowship DAMD 17-00-1-0269.

§ To whom correspondence should be addressed: Dept. of Oncology/Lombardi Cancer Center, Georgetown University Medical Ctr., 3970 Reservoir Rd. NW, Washington, DC 20007. Tel.: 202-687-4304; Fax: 202-687-7505; E-mail address: lincy@gunet.georgetown.edu.

[1] The abbreviations used are: ECM, extracellular matrix; HGF, hepatocyte growth factor; SF, scattering factor; HAI-1, hepatocyte growth factor activator inhibitor-1; LDL, low density lipoprotein; MT-SP1, membrane-type serine protease 1; uPA, urokinase plasminogen activator; pro-uPA, single-chain form of human uPA; MDCK, Madin-Darby canine kidney; PAGE, polyacrylamide gel electrophoresis; FCS, fetal calf serum; PBS, phosphate-buffered saline; AMC, 7-amino-4-methylcoumarin; *N-t*-Boc, *N-tert*-butoxycarbonyl.

Biotechnologies (Calne, Wilshire, UK). Monoclonal anti-human Met antibodies (clones DL-21 and DL-24) were purchased from Upstate Biotechnology Inc. (Lake Placid, NY). The recombinant anti-phosphotyrosine antibody (RC20: HRPO) was from Transduction Laboratory (Lexington, KY).

Cell Lines and Protein Substrates—Human lung carcinoma cell line A549 was from the ATCC. Madin-Darby canine kidney (MDCK II) epithelial cell lines and the single-chain form HGF protein were the generous gifts from Dr. George Vande Woude (Van Andel Research Institute, Grand Rapids, MI). The single-chain form of human urokinase plasminogen activator (pro-uPA) was purchased from American Diagnostics Inc. (Greenwich, CT). Plasminogen, plasmin, and fluorescent substrate peptides *N-tert*-butoxycarbonyl (*N-t*-Boc)-Gln-Ala-Arg-7-Amido-4-methylcoumarin (AMC), *N-t*-Boc-Gln-Gly-Arg-AMC, *N-t*-Boc-Leu-Gly-Arg-AMC, *N-t*-Boc-γ-benzyl (Bz)-Glu-Gly-Arg-AMC, *N-t*-Boc-γ-Bz-Glu-Ala-Arg-AMC, *N*-succinyl (Suc)-Ala-Phe-Lys-AMC, *N*-Suc-Leu-Leu-Val-Tyr-AMC, Suc-Ala-Ala-Pro-Phe-AMC, and Suc-Ala-Ala-Ala-AMC were from Sigma Chemical Co. (St. Louis, MO).

Purification of Active Matriptase —The 70-kDa active matriptase and its endogenous inhibitor HAI-1 were purified from human milk by immunoaffinity chromatography and maintained in their uncomplexed status in glycine buffer, pH 2.4, as described previously (7). Matriptase and HAI-1 were further separated by 10% SDS-polyacrylamide gel electrophoresis (SDS-PAGE), and the gel was stained using a zinc stain kit (Bio-Rad, Hercules, CA). The 70-kDa active matriptase was sliced out and then eluted from the gel using Electro-Eluter (Bio-Rad) under non-denatured conditions (Tris-glysine buffer, pH 8.3). Alternatively, the active matriptase was purified by a high pressure liquid chromatography C18 column (Vydac VHPB5). The mobile phase was A = 0.1% trifluoroacetic acid in water; B = 95% acetonitrile in water containing 0.1% trifluoroacetic acid. The flow rate was 1 ml/min. The gradient was set from 0 to 60% B over 8 min and at 60% B for next 7 min. The 70-kDa active matriptase was composed of two major and one minor cleaved products of the membrane-bound matriptase. The cleavage sites of the two major bands were identified to be Lys^{189}-Ser^{190} and Lys^{204}-Thr^{205} (the numbering of amino acid residues are based on the updated cDNA sequence of matriptase). The 70-kDa, active matriptase contains two CUB domains, four LDL receptor domains, and the serine protease domain but lacks the transmembrane domain.

Determination of Matriptase Activity—The enzyme activity of matriptase was measured at room temperature in a reaction buffer containing 100 mM Tris-HCl (pH 8.5) and 100 μg/ml bovine serum albumin, using a fluorescent peptide as substrate. In brief, 10 μl of enzyme solution and 10 μl of peptide substrate were added to a cuvette containing 180 μl of the reaction buffer. The mixture was mixed well, placed back into a fluorescent spectrophotometer (Hitachi F4500), and the release of fluorescence resulting from hydrolysis of the peptide substrate was recorded with excitation at 360 nm and emission at 480 nm.

Determination of Kinetic Parameter—Substrate concentration *versus* initial reaction velocity were analyzed by the Michaelis-Menten equation and plotted using SigmaPlot software. Double reciprocal (Lineweaver-Burk) plots thus derived were used to determine V_{max} and K_m values.

Cleavage of Protein Substrates—Single chain HGF protein, plasminogen, or pro-uPA was incubated with various amounts of matriptase in 100 mM Tris-HCl (pH 8.5) overnight at room temperature. Incubation was stopped by boiling the mixture in SDS sample buffer. The cleaved products were then separated on SDS-PAGE and analyzed by Western blot hybridization or by silver stain.

Scattering Assay—The MDCK II cell line was maintained at 37 °C in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum (FCS). The scatter assay was carried out in the 96-well culture plate (11). To each well was added 150 μl of Dulbecco's modified Eagle's medium supplemented with 5% FCS and leupeptin at 100 μg/ml; HGF or matriptase-cleaved HGF was added into one well and serial 2-fold dilutions were made with sequential 150-μl aliquots of medium transferred from well to well. About 3000 MDCK II cells in 150 μl of medium were added to each well, and the plate was placed in a 37 °C incubator for 20–24 h. Media were removed, and cells were fixed and stained for 15 min with 5% crystal violet in 50% methanol. Cell scattering (spreading and dispersion of epithelial colonies) was examined under light microscopy.

c-Met Phosphorylation Detection—A549 cells were grown confluent in RPMI medium supplemented with 10% FCS. After 3-hour serum starvation, cells were incubated 5 min at 37 °C with 450 ng/ml HGF or matriptase-cleaved HGF in RPMI medium supplemented with 5% FCS.

TABLE I

Kinetic parameters of matriptase for various peptide substrates

10 μl of matripase solution and 10 μl of peptide substrate were added to a cuvette containing 180 μl of reaction buffer, and the release of fluorescence resulting from hydrolysis of the peptide was recorded in a fluorescence spectrophotometer at room temperature with excitation at 360 nm and emission at 480 nm. K_m and V_{max} were determined by double-reciprocal plot derived from the Michaelis-Menten equation.

Peptide substrate	P4-P3-P2-P1-AMC	K_m	V_{max}
		μM	nM AMC/min
1	Gln-Ala-Arg[a]	4.89	654
2	Glu-Ala-Arg[a]	3.81	76.3
3	Leu-Gly-Arg[a]	13.6	309
4	Gln-Gly-Arg[a]	33.5	528
5	Glu-Gly-Arg	47.5	170
6	Ala-Phe-Lys	69.9	524
7	Leu-Leu-Val-Tyr[a]	-[b]	-[b]
8	Ala-Ala-Pro-Phe[a]	-[b]	-[b]
9	Ala-Ala-Ala[a]	-[b]	-[b]

[a] Gln-Ala-Arg is a standard substrate for trypsin, Glu-Ala-Arg is a substrate for factor XIa, Leu-Gly-Arg is a substrate for uPA, Gln-Gly-Arg is a substrate for XIIa, Leu-Leu-Val-Tyr and Ala-Ala-Pro-Phe are substrates for chymotrypsin, and Ala-Ala-Ala is the substrate for elastase.

[b] No cleavage activity was detected with these substrates at a concentration of 200 μM.

As appropriate, leupeptin was included in the medium at 100 μg/ml. Media were removed, and cells were rinsed with 1× phosphate-buffered saline (PBS) and collected by centrifugation following trypsinization. After washing one more time with 1× PBS, cell pellets were frozen in dry ice. The frozen cell pellets were either stored at −80 °C for later extraction or immediately extracted as described below. Cells were thawed on ice, extracted by suspension in buffer containing 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 150 mM sodium fluoride, 10 mM sodium pyrophosphate, 10 mM sodium orthovanadate, 100 μg/ml phenylmethylsulfonyl fluoride, 10 μg/ml leupeptin, 10 μg/ml aprotinin, and 1% Triton X-100. Extracts were clarified by centrifugation for 15 min at 12,000 × *g* in a microcentrifuge, and the protein concentration was determined by BCA protein assay kit (Pierce, Rockford, IL) using bovine serum albumin as standard. About 2 mg of protein of extracts was immunoprecipitated using anti-c-Met antibody and Pansorbin (Calbiochem, La Jolla, CA). The protein-antibody-Pansorbin immunocomplex was collected by centrifugation, washed twice with extraction buffer then with 1× PBS, and then dissociated by boiling in SDS-sample buffer. Pansorbin was removed by centrifugation, the supernatant fractions were subjected to 8% SDS-PAGE, and the proteins were detected by Western immunoblot using anti-phosphotyrosin antibody. The same immunoblots were then stripped in 100 mM 2-mercaptoethonal, 2% SDS, and 62.5 mM Tris-HCl (pH 7.6) for 30 min at 50 °C and reprobed with anti-c-Met antibody.

RESULTS

Matriptase Selectively Cleaves Peptide after an Arg or Lys Residue—In our previous study, the trypsin-like activity of matriptase was suggested by the observations that 1) an Asp residue positioned at the bottom of a substrate binding pocket of the serine protease domain of matriptase and 2) matriptase is able to cleave various synthetic substrates containing Arg or Lys as P1 sites (6). In the current study, to investigate in detail the substrate specificity of matriptase, we measured the K_m and V_{max} of matriptase for a variety of protease substrate peptides. Table I shows that the most reactive peptide substrates for matriptase are *N-t*-Boc-Gln-Ala-Arg-AMC, with a K_m of 4.89 μM, and N-t-Boc-Bz-Glu-Ala-Arg-AMC, with a K_m of 3.81 μM. *N-t*-Boc-Gln-Ala-Arg-AMC and N-t-Boc-Bz-Glu-Ala-Arg-AMC were reported to be good substrates for bovine trypsin and human factor XIa, respectively (12). No released fluorescence was detected from the substrates for chymotrypsin or elastase (Table I, peptide substrates 7, 8, and 9). Matriptase

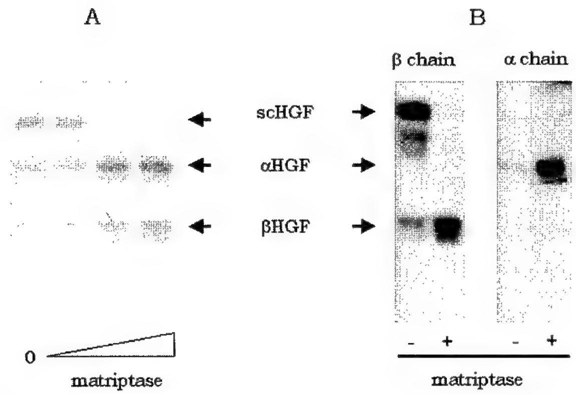


FIG. 1. **Matriptase converts single-chain HGF into smaller fragments, which can be recognized by anti-α-chain HGF and β-chain HGF antibodies.** *A*, protein staining. Shown are the silver-stained protein patterns of HGF incubated overnight without (*0*) or with increasing amount of matriptase. *B*, Western immunoblot. HGF incubated overnight without (−) or with (+) matriptase was immunoblotted with anti-βHGF (*β chain*) or anti-αHGF (*α chain*).

appears to prefer to bind to peptides containing small side-chain amino acids, such as Ala and Gly, at P2 site (Table I, peptide substrates 1–5). Peptides containing P2 Ala are better substrates for matriptase than peptides containing P2 Gly (compare peptides 1 and 2 with peptides 3–5). The binding affinity of matriptase to the former is about 30-fold higher than that to the latter. Interestingly, a change from Gln to Glu at the P3 site significant reduces the V_{max} (compare peptide 1 with 2) without causing a significant change to the K_m.

Considering its trypsin-like activity and its presentation on the surfaces of epithelial cells and cancer cells, matriptase could serve as a cell surface activator for other secreted proteases and growth factors. Particularly relevant substrates could include those factors that are produced by stromal cells, that function on epithelial cells, and that require proteolytic cleavage at Arg or Lys for their activation. We, therefore, set out to examine whether matriptase can activate three such substrates, HGF/SF, plasminogen, and pro-uPA, in addition to characterizing the synthetic substrates for matriptase.

Matriptase Is Able to Activate HGF/SF—HGF/SF is secreted as an inactive, single chain precursor by stromal cells, and it is activated by proteolytic conversion to the two-chain form factor by cleavage at Arg^{495} (13) in the extracellular environment. This cleavage is required for HGF/SF to activate its cell surface receptor c-Met. To determine if HGF/SF can be cleaved by matriptase, inactive HGF/SF, purified from the condition medium of fibroblast cells cultured in the absence of serum (a generous gift from Dr. Vande Woude, Van Andel Research Institute) was utilized in the following experiments. Fig. 1*A* shows that this HGF/SF preparation is primarily composed of the single-chain form protein, with an apparent size of about 97 kDa on SDS-PAGE (*left lane*). Because it was reported previously that there was a limited contamination of cleaved form HGF in preparations of latent form HGF (14, 15), we wished to establish whether there was contamination in our preparation of latent HGF. As expected, we also observed minor species of 64- and 33-kDa apparent molecular mass, corresponding to the expected sizes of activation cleaved α- and β-chains of HGF/SF, respectively. After incubation with the active form matriptase, the amount of 64- and 33-kDa molecules increased at the expense of a 97-kDa molecule. The lowest amount of matriptase allowing observable HGF cleavage was at a molar ratio of approximately 500:1 for HGF:matriptase. Western immunoblot analysis using monoclonal antibodies to α- and β-chain HGF showed that the 64-kDa species contains αHGF and the 33-kDa species contains βHGF (Fig. 1*B*). The contaminating 64- and 33-kDa species in the latent HGF sample in the absence of incubation with matriptase were also detected with the respective antibodies (Fig. 1*B*). The anti-αHGF antibody apparently does not recognize the single-chain form HGF/SF (Fig. 1*B*). Thus, matriptase can convert the single chain HGF/SF into fragments containing α- and β-HGF. We scanned the silver-stained protein gel and the x-ray film of this Western immunoblot by densitometry and estimated the relative amounts of the full-length, α-, and β-HGF (data not shown) using ImageQuant software (Molecular Dynamics). On average, we found that areas under the α- and β-chains were 25% and 30%, respectively, of the total stained HGF. The α- and β-chains in the original HGF preparation were 25–30% of their respective values in the matriptase-cleaved sample, as determined by both silver-staining and Western immunoblotting. Therefore, we estimated that approximately 25% of the HGF in our preparation of latent HGF is in the cleaved (activated) form.

To further investigate if the cleavage of HGF/SF by matriptase corresponds to a process of activation, we performed a well-characterized assay specific to HGF activity, the MDCK cell scatter assays (2, 11). The single-chain HGF/SF was first incubated with matriptase to allow its complete cleavage; the cleaved products were than diluted in the medium and incubated with MDCK II cells for 20–24 h at 37 °C. To eliminate the possibility of activation of HGF/SF by other activators in the serum, leupeptin was included in every scattering assay (16). It was previously observed that leupeptin can also inhibit matriptase enzyme activity (4). Fig. 2 shows the images of plates of MDCK II cells after incubation with various dilutions of untreated (HGF) or matriptase-treated (HGF/MTP) HGF/SF. In the control sample (CRL), cells aggregated into tight clusters after 20–24 h culture. In the presence of 1:4096 dilution of HGF/MTP, cell islands were more scattered, and there was a high proportion of single cells. In contrast, at the same dilution of HGF, cells were not significantly different from the control cells. The scatter morphology was more profound with progressively lower dilutions of HGF/MTP. At 1:512 dilution of HGF/MTP, cells completely separated from each other and many of them exhibited elongated, fibroblast-like shapes. On the other hand, scattering morphology induced by the untreated HGF appeared at much lower dilution (Fig. 2, HGF, 1:512). For the cells incubated with untreated HGF, we consistently observed the first sign of scatter at 1:1024 dilution, compared with 1:4096 dilution for the matriptase-cleaved HGF. This difference in the scattering activity comparing matriptase-cleaved and uncleaved HGF is very close to the difference observed in the protein level demonstrated in Fig. 1. Thus the baseline scattering activity in samples treated with uncleaved HGF apparently was derived from the contaminating active HGF in the HGF/SF preparation (Fig. 1). Cells grown in the presence of matriptase alone retained the same morphology as those in the control (data not shown). Cells incubated with untreated HGF/SF in the absence of leupeptin were examined as a positive control for HGF activation by serum; the scattering activity of HGF/SF in these samples was as expected (data not shown). These observations showed that the scattering induced by HGF/MTP is indeed enhanced by the matriptase activation of HGF.

HGF functions through binding to its cell surface receptor, c-Met. Upon binding to HGF, c-Met is activated following phosphorylation at tyrosine residues. In Fig. 3, c-Met activation was examined in the A549 human lung carcinoma cell line. Like the experiments for Fig. 2, leupeptin was also included in these treatments. Tyrosine-phosphorylated c-Met was observed in cells incubated with matriptase-cleaved HGF (HGF/MTP) and with untreated HGF (HGF), but not in cells incubated in the

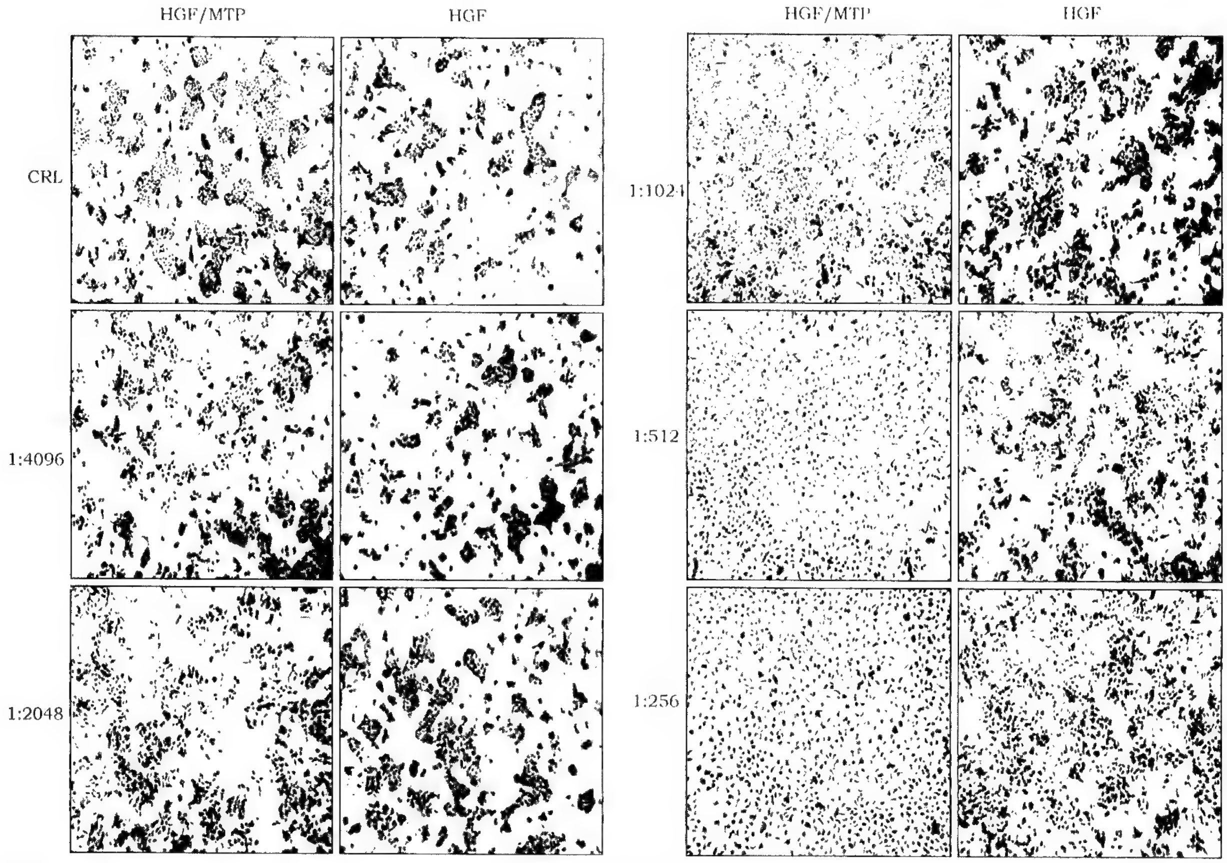


FIG. 2. **Matriptase-cleaved HGF stimulates scattering on MDCK cells.** Plates show wells in the 96-well plates with about 3000 MDCK cells after 20 h in the absence (*CRL*) or presence of matriptase-cleaved HGF (*HGF/MTP*) or untreated HGF (*HGF*) at the dilutions shown. Leupeptin was included in every culture.

absence of HGF (No HGF). The tyrosine-phosphorylated c-Met in the cells treated with uncleaved HGF is about one-fifth of that in the cells treated with matriptase-cleaved HGF. Again, consistent with a limited degree of contamination of cleaved HGF in our latent HGF preparation. Phosphotyrosine detected in c-Met of cells incubated with untreated HGF appears to be caused by the residual active HGF contamination in the preparation. Leupeptin did not affect the total c-Met expression, the c-Met phosphorylation, nor the total pattern of tyrosine phosphorylation (data not shown).

Plasminogen shares high homology with HGF, and its activation also requires a cleavage at Arg. Therefore, it seemed likely that plasminogen would be a substrate of matriptase as well. However, matriptase failed to cleave plasminogen. As shown in Fig. 4, plasminogen remained as a 94-kDa single-chain form, even in the presence of matriptase at a concentration 8-fold higher than that required for cleavage of HGF.

Matriptase Could Function as an Initiator of Matrix-degrading Protease Cascade—Plasmin has long been regarded as the enzyme that converts pro-uPA to active uPA. However, the level of active uPA is not reduced in the urine of mice bearing a targeted disruption of the plasminogen gene (17), suggesting the existence of plasmin-independent pro-uPA activation. Plasma kallikrein (18), trypsin-like proteases from human ovarian tumors (19), a T cell-associated serine protease (20), cathepsins B and L (21, 22), nerve growth factor γ (23), human mast cell tryptase (24), and prostate-specific antigen (25) have also been reported to activate pro-uPA. However, the relevance of these studies of pro-uPA activation *in vitro* is uncertain for understanding roles of these enzymes *in vivo*. During the preparation of this manuscript, activation of pro-uPA was reported by the recombinant serine protease domain of matriptase/MT-

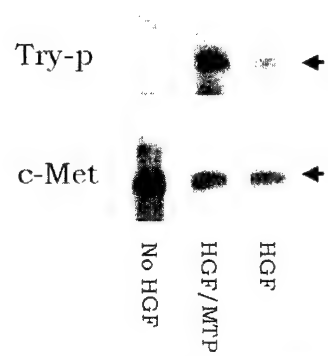


FIG. 3. **Matriptase-cleaved HGF stimulates c-Met tyrosine phosphorylation.** A549 cells were treated in the absence of HGF (*No*), matriptase-cleaved HGF (*HGF/MTP*), or HGF alone (*HGF*) for 5 min at 37 °C. Equal amounts of lysed cell protein were immunoprecipitated with anti-c-Met antibody followed by immunodetection using anti-phosphotyrosin antibody (*Tyr-p*) as described under "Materials and Methods." To control for the amounts of c-Met in each sample, immunoblots were then stripped and detected with anti-c-Met antibody (*c-Met*). The *arrow* points to the 145-kDa β-chain of c-Met under reducing conditions.

SP1 (26). This observation is consistent with our studies using the purified, 70-kDa active matriptase containing both CUB and LDL domains. Fig. 5*A* showed that, after incubation with matriptase, the 55-kDa single-chain pro-uPA was converted into smaller fragments. One of these cleavage products clearly appeared on the protein gel as the 33-kDa molecule, which resembles the size of the active uPA protease (Fig. 5*A*). The cleaved product exhibited enzymatic activity toward the fluorescent peptide substrate, *N-t*-Boc-Leu-Gly-Arg-AMC, for uPA (Fig. 5*B*, compare the *closed circles* with *closed triangles*). This

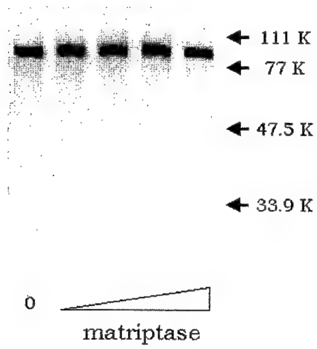


FIG. 4. **Plasminogen is not a substrate for matriptase.** Shown are the silver-stained protein patterns of plasminogen incubated overnight without (*0*) or with increasing amount of matriptase. The highest amount of matriptase used is 8-fold of the lowest amount of matriptase that cleaves HGF.

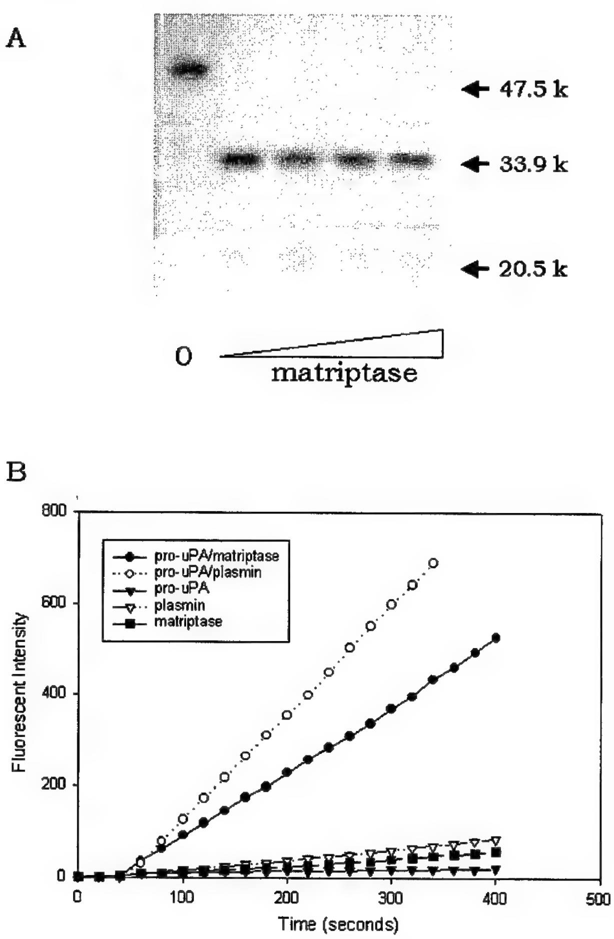


FIG. 5. **Pro-uPA is activated by matriptase cleavage.** *A*, single-chain pro-uPA is converted into two-chain form uPA by matriptase. Pro-uPA was incubated overnight with active matriptase in the absence (*0*) or presence of increasing amount of matriptase. The cleaved products were analyzed by electrophoresis followed by silver-staining. *B*, matriptase cleavage of pro-uPA generates an active protease. Pro-uPA was either incubated for 30 min with matriptase (*closed circles*) or incubated 1 min with plasmin (*opened circles*) prior to the assay. Plasmin (*opened triangles*), uPA (*closed triangles*), and matriptase (*closed squares*) all exhibit low activity.

activity was not derived from matriptase, because matriptase alone only exhibited negligible background activity (Fig. 5*B*, *closed squares*). The same amounts of pro-uPA produced a similar activity after cleavage by plasmin (Fig. 5*B*, *opened circles*). These results suggest that matriptase itself is able to activate pro-uPA and that the CUB domains and LDL receptor domains of matriptase do not interfere with its activation activity.

DISCUSSION

By using the 70-kDa, active matriptase isolated from human milk, we report in this study that matriptase cleaves and converts HGF into a biologically functional factor that can induce c-Met activation and stimulate epithelial cell scattering. In addition, we also noted that matriptase can activate pro-uPA but not plasminogen. These results further support our working hypothesis that matriptase is an upstream regulator of cellular migration and extracellular matrix degradation. Most significantly, these results reveal a novel mechanism in the control of tissue remodeling that involves an upstream epithelial membrane activator and downstream stromal effectors.

Tissue remodeling is a process observed both in physiological and pathologic processes. Two essential changes occur during these processes: 1) an epithelial-mesenchymal transition transforms relatively rigid epithelial cells to the more mobile migratory mesenchymal cells; 2) extracellular matrix degradation opens pathways for the migrating cells. HGF/SF is a potent inducer of epithelial-mesenchymal transition (27); engagement of HGF to its epithelial receptor c-Met triggers various intracellular signaling pathways. HGF is secreted as an inactive precursor by stromal cells, and it is proteolytically activated in the extracellular environment (13). Therefore, activation of HGF/SF needs to occur in the close vicinity of the epithelial cells. uPA and the zinc-dependent metalloproteinases have been proposed to be responsible for the majority of proteolysis of pericellular proteins (28). However, both systems are largely synthesized by the stromal cells and require indirect mechanisms for their recruitment and activation on the surfaces of epithelial cells. Thus, an epithelial-derived protease, like matriptase, could provide a missing link in this process.

Matriptase appears to have selectivity for its macromolecular substrates. In our experiments, matriptase did not cleave plasminogen, despite the high sequence homology between plasminogen and HGF. This selectivity was also reported by others utilizing the serine protease domain of matriptase (26). In the same report, it was also shown that matriptase/MT-SP1 has selectivity for a basic residue at the P3 or P4 site. The sequence at the activation cleavage site of HGF and plasminogen is P4-(Lys)-P3-(Gln)-P2-(Gly)-P1-(Arg) and P4-(Cys)-P3-(Pro)-P2-(Leu)-P1-(Arg), respectively. Thus, the lack of a P3 or P4 basic residue in plasminogen might contribute to its lack of cleavage by matriptase. Alternatively, there might be conformational differences between HGF and plasminogen. The kringle domains of plasminogen contain a Lys-binding site that serves to mediate its localization to fibrin and to cellular surfaces. Plasminogen circulates in the blood in a globular and closed conformation; upon binding to the surface, it shifts to an extended and opened conformation. This conformation change promotes its recognition by its activator and its rapid conversion to plasmin. The kringle domains on HGF/SF also contribute to its binding to its receptor c-Met. However, cleavage activation of HGF/SF does not depend on its binding to c-Met. It is possible that the single chain form HGF/SF exhibits a more open conformation than does plasminogen and that matriptase can distinguish this subtle structural difference.

Acknowledgments—We thank Dr. George Vande Woude (Van Andel Research Institute, Grand Rapids, MI) for providing the single-chain form HGF and the MDCK II cell line. We also thank Marianne Oskarsson (National Cancer Institute, NIH, Frederick, MD) for valuable suggestions on our scattering assays. We also thank Dr. Radoslav Goldman for the purification of matriptase using high pressure liquid chromatography.

REFERENCES

1. Dano, K., Andreasen, P. A., Grondahl-Hansen, J., Kristensen, P., Nielsen, L. S., and Skriver, L. (1985) *Adv. Cancer Res.* **44,** 139–266
2. Stoker, M., Gherardi, E., Perryman, M., and Gray, J. (1987) *Nature* **327,** 239–242
3. Dano, K., Romer, J., Nielsen, B. S., Bjorn, S., Pyke, C., Rygaard, J., and Lund, L. R. (1999) *APMIS* **107,** 120–127
4. Shi, Y. E., Torri, J., Yieh, L., Wellstein, A., Lippman, M. E., and Dickson, R. B. (1993) *Cancer Res.* **53,** 1409–1415
5. Lin, C. Y., Wang, J. K., Torri, J., Dou, L., Sang, Q. X. A., and Dickson, R. B. (1997) *J. Biol. Chem.* **272,** 9147–9152
6. Lin, C. Y., Anders, J., Johnson, M., Sang, Q. A., and Dickson, R. B. (1999) *J. Biol. Chem.* **274,** 18231–18236
7. Lin, C. Y., Anders, J., Johnson, M., and Dickson, R. B. (1999) *J. Biol. Chem.* **274,** 18237–18242
8. Takeuchi, T., Shuman, M. A., and Craik, C. S. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 11054–11061
9. Kim, M. G., Chen, C., Lyu, M. S., Cho, E. G., Park, D., Kozak, C., and Schwartz, R. H. (1999) *Immunogenetics* **49,** 420–428
10. Shimomura, T., Denda, K., Kitamura, A., Kawaguchi, T., Kito, M., Kondo, J., Kagaya, S., Qin, L., Takata, H., Miyazawa, K., and Kitamura, N. (1997) *J. Biol. Chem.* **272,** 6370–6376
11. Stoker, M., and Perryman, M. (85 A. D.) *J. Cell Sci.* **77,** 209–223
12. Kawabata, S., Miura, T., Morita, T., Kato, H., Fujikawa, K., Iwanaga, S., Takada, K., Kimura, T., and Sakakibara, S. (1988) *Eur. J. Biochem.* **172,** 17–25
13. Naka, D., Ishii, T., Yoshiyama, Y., Miyazawa, K., Hara, H., Hishida, T., and Kidamura, N. (1992) *J. Biol. Chem.* **267,** 20114–20119
14. Weidner, K. M., Behrens, J., Vandekerckhove, J., and Birchmeier, W. (1990) *J. Cell Biol.* **111,** 2097–2108
15. Rong, S., Oskarsson, M., Faletto, D., Tsarfaty, I., Resau, J. H., Nakamura, T., Rosen, E., Hopkins, R. F., 3rd, and Vande Woude, G. F. (1993) *Cell Growth Differ.* **4,** 563–569
16. Naldini, L., Tamagnone, L., Vigna, E., Sachs, M., Hartmann, G., Birchmeier, W., Daikuhara, Y., Tsubouchi, H., Blasi, F., and Comoglio, P. M. (1992) *EMBO J.* **11,** 4825–4833
17. Bugge, T. H., Flick, M. J., Daugherty, C. C., and Degen, J. L. (1995) *Genes Dev.* **9,** 794–807
18. Ichinose, A., Fujikawa, K., and Suyama, T. (1986) *J. Biol. Chem.* **261,** 3486–3489
19. Koivunen, E., Huhtala, M. L., and Stenman, U. H. (1989) *J. Biol. Chem.* **264,** 14095–14099
20. Brunner, G., Vettel, U., Jobstmann, S., Kramer, M. D., and Schirrmacher, V. (1992) *Blood* **79,** 2099–2106
21. Kobayashi, H., Schmitt, M., Goretzki, L., Chucholowski, N., Calvete, J., Kramer, M., Gunzler, W. A., Janicke, F., and Graeff, H. (1991) *J. Biol. Chem.* **266,** 5147–5152
22. Goretzki, L., Schmitt, M., Mann, K., Calvete, J., Chucholowski, N., Kramer, M., Gunzler, W. A., Janicke, F., and Graeff, H. (1992) *FEBS Lett.* **297,** 112–118
23. Wolf, B. B., Vasudevan, J., Henkin, J., and Gonias, S. L. (1993) *J. Biol. Chem.* **268,** 16327–16331
24. Stack, M. S., and Johnson, D. A. (1994) *J. Biol. Chem.* **269,** 9416–9419
25. Yoshida, E., Ohmura, S., Sugiki, M., Maruyama, M., and Mihara, H. (1995) *Int. J. Cancer* **63,** 863–865
26. Takeuchi, T., Harris, J. L., Huang, W., Yan, K. W., Coughlin, S. R., and Craik, C. S. (2000) *J. Biol. Chem.* **275,** 26333–26342
27. Stewart, F. (1996) *Rev. Reprod.* **1,** 144–148
28. Mignatti, P., and Rifkin, D. B. (1993) *Physiol. Rev.* **73,** 161–195

Structure-Based Approach for the Discovery of Bis-benzamidines as Novel Inhibitors of Matriptase

Istvan J. Enyedy, Sheau-Ling Lee, Angera H. Kuo, Robert B. Dickson, Chen-Yong Lin, and Shaomeng Wang

Structural Biology and Cancer Drug Discovery Program, Lombardi Cancer Center, and Department of Oncology, Georgetown University Medical Center, 3900 Reservoir Road, Washington D.C. 20007

JOURNAL OF

MEDICINAL CHEMISTRY®

Reprinted from

Volume 44, Number 9, Pages 1349–1355

Structure-Based Approach for the Discovery of Bis-benzamidines as Novel Inhibitors of Matriptase

Istvan J. Enyedy,† Sheau-Ling Lee,† Angera H. Kuo, Robert B. Dickson, Chen-Yong Lin, and Shaomeng Wang*

Structural Biology and Cancer Drug Discovery Program, Lombardi Cancer Center, and Department of Oncology, Georgetown University Medical Center, 3900 Reservoir Road, Washington D.C. 20007

Received September 8, 2000

Matriptase, a trypsin-like serine protease, which may be involved in tissue remodeling, cancer invasion, and metastasis. Potent and selective matriptase inhibitors not only would be useful pharmacological tools for further elucidation of the role of matriptase in these processes but also could have therapeutic potential for the treatment and/or prevention of cancers. We report herein the structure-based approach for the discovery of bis-benzamidines as a novel class of potent matriptase inhibitors. The lead compound, hexamidine (**1**), inhibits not only the proteolytic activity of matriptase, (K_i = 924 nM) but also of thrombin K_i = 224 nM). By testing several available analogues, we identified a new analogue (**7**) that has a K_i = 208 nM against matriptase and has only weak inhibitory activity against thrombin (K_i = 2670 nM), thus displaying a 13-fold selectivity toward matriptase. Our results demonstrated that structure-based database screening is effective in the discovery of matriptase inhibitors and that bis-benzamidines represent a class of promising matriptase inhibitors that can be used for further drug design studies. Finally, our study suggested that there is sufficient structural differences between matriptase and its closely related serine proteases, such as thrombin, for the design of potent and selective matriptase inhibitors.

Introduction

Local invasion and metastasis of cancers have been proposed to require imbalanced or unregulated expression of proteases, such as metalloproteases and urokinase-type plasminogen activator (uPA), at invading edges of carcinoma cells.[1–5] In recent years, increasing efforts have been applied to the development of potent and selective inhibitors of these proteases as potential anticancer therapeutic agents.[6–8] Indeed, a number of metalloprotease inhibitors are now in clinical trial for the treatment of cancer.[7]

We have recently characterized a novel, integral membrane serine protease, matriptase (GenBank accession number AF118224), and its cognate inhibitor HAI-1 (hepatocyte growth factor activator inhibitor 1).[9–11] In contrast to most other protease–inhibitor systems, both matriptase and HAI-1 are selectively expressed by cultured breast epithelial cells and cancer cells, but not by fibroblasts or fibrosarcoma cells. In addition to the C-terminal serine protease domain, the N-terminal noncatalytic region of matriptase contains two tandem repeats of a CUB (*C*1r/s, *U*egf, and *B*one morphogenetic protein-1) domain and four tandem repeats of a low-density lipoprotein receptor domain that are likely to be involved in protein–protein interaction.[10] Immunofluorescent staining of cultured breast cancer cells demonstrated that matriptase is concentrated on the cell peripheries at pseudopodia and on membrane ruffles in spreading cells.[9] Of particular interest, matriptase has recently been shown to activate hepatocyte growth factor and single-chain urokinase plasminogen activator.[5,12] Recently, a large-peptide inhibitor of matriptase, ecotin, has been shown to retard the growth of PC-3 prostate cancer tumors in nude mice, further suggesting that matriptase may play a role in cancer progression.[13] Taken together, these data suggest that matriptase may be a central regulator of cell migration and cancer invasion and may provide a novel cancer-associated protease target for design of new anticancer drugs.

We are interested in the discovery and development of potent and selective small-molecule inhibitors of matriptase. A potent, selective, nonpeptide, and drug-like small-molecule matriptase inhibitor not only will be useful as a pharmacological tool to further elucidate the biological functions of matriptase but also may have the therapeutic potential for treatment of cancer through stopping invasion and metastasis. Although large-peptide matriptase inhibitors, such as ecotin, have been reported, drug-like, small-molecule matriptase inhibitors are currently not available. Herein, we report our discovery of bis-benzamidines as a class of matriptase inhibitors through structure-based database search.

Results and Discussion

Homology Modeling. The X-ray structure of human thrombin, entry 1hxe (Figure 1A) from the Protein Data Bank,[14,15] was used as template for building the 3D structure of the protease domain of matriptase using homology modeling. It was shown that when the sequence identity/similarity between the modeled protein and the template is between 30% and 40%, the expected main-chain rms deviation between the modeled and the experimental structures for the protein is about 1.5 Å

* To whom correspondence and requests of reprints should be addressed. Phone: (202)-687-2028. Fax: (202)-687-4032. E-mail: wangs@giccs.georgetown.edu.

† These authors contributed equally to this work.

10.1021/jm000395x CCC: $20.00
Published on Web 03/29/2001

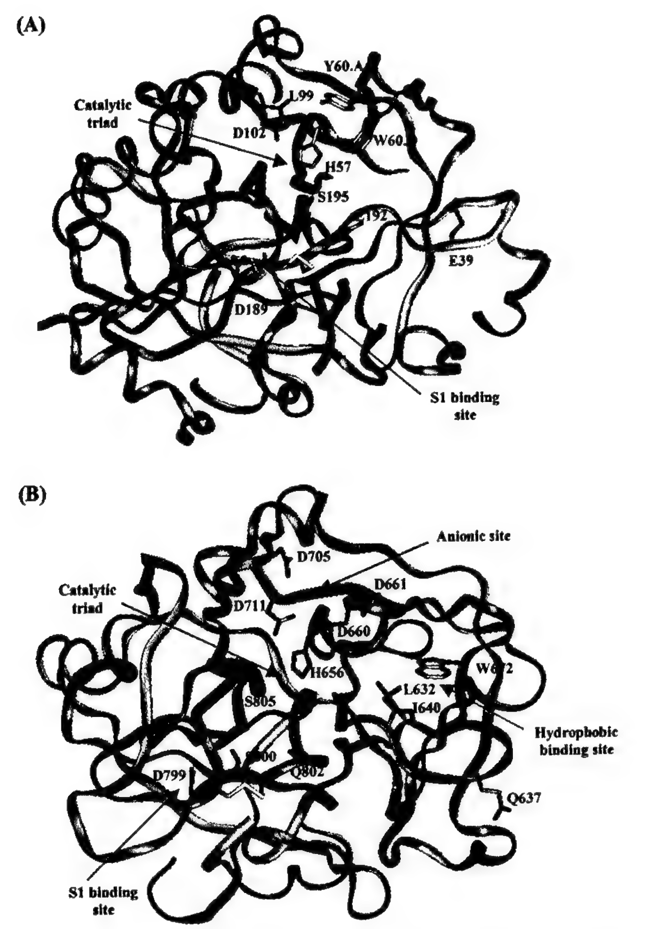


Figure 1. A: Active site residues in thrombin, entry 1hxe in the Protein Data Bank. B: Active site of matriptase, as obtained after homology modeling and refinement using MD simulation in water.

for 80% of residues.[16,17] Since the sequence identity and similarity between thrombin (Figure 1A) and matriptase are 34% and 53%, respectively, and both enzymes belong to the same protease family, it is expected that the 3D structure of matriptase can be modeled accurately. Figure 1B shows the modeled structure of the protease domain of matriptase. By analogy to thrombin, the serine protease domain of matriptase has a catalytic triad positioned on the surface, marked by Ser805, His656, and Asp711 corresponding to Ser195, His57, and Asp102, respectively in thrombin. Consistent with the observation that matriptase prefers substrates with an Arg or Lys as P1 residue,[10,12] a negatively charged residue, Asp799, is located at the bottom of the S1 binding site (Figure 1B). This residue corresponds to Asp189 in thrombin. Ser800 in the S1 binding site in matriptase differs from the corresponding Ala190 in thrombin. Close to the S1 site, Gln802 and Gln637 in matriptase correspond to charged Glu192 and Glu39, respectively, in thrombin (Figure 1). Anionic site residues Asp705, Asp660, and Asp661 in matriptase differ from the corresponding Trp96, Tyr60.A, and Trp60.D, respectively in thrombin. This shows that this site in matriptase is charged while it is neutral in thrombin. A putative hydrophobic S1′ binding site in matriptase is marked by Leu632, Ile640, and Trp672, as shown in Figure 1B, that is similar to the hydrophobic pocket formed by Leu33, Leu41, and Leu64 in thrombin, Figure 1A.

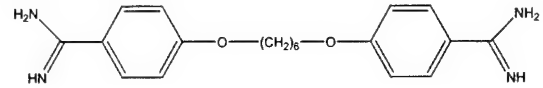
Table 1. Initial Screening of Compound Inhibitors for Matriptase

behavior	number of compounds
over 95% inhibition	15
90–94% inhibition	4
70–89% inhibition	15
40–69% inhibition	13
below 39% inhibition	17
high absorbency	3
increase activity	3

Structure-Based 3D Database Screening. The refined structure of matriptase, obtained from molecular dynamics (MD) simulation, was used for structure-based screening of the NCI database.[18] Since the S1 site is considered to be the primary binding site in serine proteases, it is likely to be a good target site for inhibitor design.[6,19] In addition, two other putative binding sites, the anionic site and the hydrophobic S1′ site, were included in the docking site used for 3D database searching with the program DOCK.[20] Ligands were scored based on the DOCK energy score computed as a sum of the electrostatic, van der Waals, and ligand conformational energy. Since the S1 site of matriptase is negatively charged, the potential inhibitor candidates that target this site should be positively charged in water under physiologic conditions for optimal interactions. Using this hypothesis a total of 69 candidate compounds were selected for testing from the best-scoring 2000 compounds based upon the DOCK program.

Inhibitory Activity Screening. Table 1 shows the results from an initial inhibitory activity screening. Each of the 69 candidate compounds was tested at 75 μM for the inhibitory activity against matriptase: 47 compounds inhibited at least 40% of the protease activity and 15 of them inhibited more than 95% of the protease activity. The 15 compounds that exhibited more than 95% inhibition were analyzed further for their K_i values, as described in the Experimental Section.

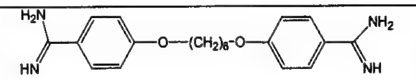
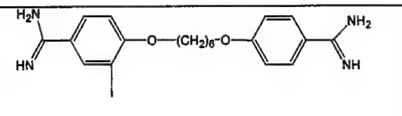
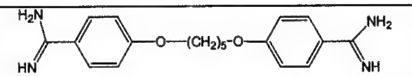
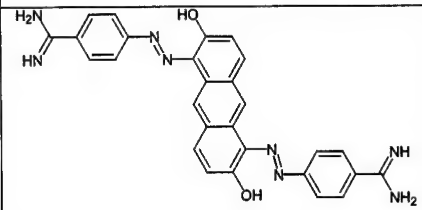
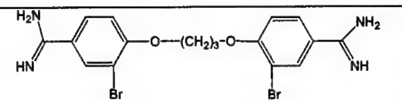
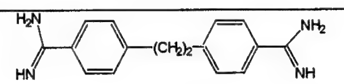
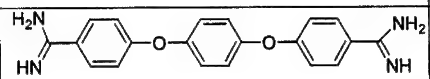
Bis-benzamidines as a Class of Potent Matriptase Inhibitors. One of the compounds that we have identified from our screening is hexamidine, compound **1**, a topical antiseptic. It inhibits matriptase with K_i = 924 nM, which makes it a good lead compound for further optimization. We therefore tested 7 closely related analogues that are available from the NCI database. Their chemical structures and K_i values for inhibition of matriptase enzymatic activity are summarized in Table 2. The K_i values of these compounds ranged from 191 nM to greater than 10 μM. Dixon plots of these inhibitors showed that they behaved as competitive inhibitors to the peptide substrate.



1 (hexamidine)

One important aspect in the design of protease inhibitors is their selectivity. For this reason compounds **1**, **2**, **5**, and **7**, with K_i values below 1 μM, were further evaluated for their selectivity against two other serine proteases, uPA and thrombin. Thrombin is a serine

Table 2. K_i Values Obtained for Bis-benzamidine Analogues of Hexamidine[a]

Cpd	Structure	K_i (nM)		
		Matriptase	uPA	Thrombin
1		924	14,400	224
2		191	1,980	796
3		1,160	N.T.	N.T.
4		4,500	N.T.	N.T.
5		535	1,570	946
6		> 10,000	N.T.	N.T.
7		208	1,950	2,670

[a] N.T. = not tested.

protease that plays a role in blood clotting, and its structure was used as the template to model the structure of matriptase. uPA is another serine protease that was proposed to play a role in cancer invasion and has a high sequence homology to matriptase.[4]

While **1** is 16-fold selective for matriptase over uPA, it is a 4-fold more selective inhibitor for thrombin over matriptase (Table 2). Compounds **1** and **2** have the same structure except for a 3-iodo substituent on one benzamidine phenyl ring. In contrast, **2** is a more selective inhibitor for matriptase; i.e., it displays 5- and 11-fold selectivity for matriptase over thrombin and uPA, respectively. This iodo substituent improved the potency for matriptase by approximately 5-fold, while it decreased the potency for thrombin by approximately 4-fold suggesting that iodo substituents on the benzamidine phenyl ring can improve the potency and selectivity of bis-benzamidines for matriptase. A structural comparison of compounds **1**, **3**, and **6** shows that the length of the linker between the two benzamidine groups plays a role for the inhibitory activity of bis-benzamidines. As the length of the linker decreases, the potency of the inhibitors decreases, which is consistent with our modeled structure of matriptase. Based upon our modeled matriptase structure, the distance between the S1 binding site and the anionic binding site is approximately 21 Å, similar to the length of **1** (20.9 Å) and **3** (19.3 Å) when they adopt a fully extended conformation. Interestingly, although compound **5** has a shorter linker than do **1** and **3**, it is approximately 2-fold more potent than **1** and **3** against matriptase. This suggests that 3,3′-dibromo substituents on the benzamidine phenyl rings improve the inhibitory potency of a compound. Compound **7** is the most rigid compound among **1**, **2**, **5**, and **7** and is also the most selective and one of the most potent inhibitors identified in this study. Despite its shorter linker **7** is as potent inhibitor as compound **2** against matriptase, but **7** has improved selectivity between matriptase and thrombin as compared to **2**. The conformational rigidity of **7** may play a role for its good potency and improved selectivity.

Docking Results. To gain a better understanding of the interactions between bis-benzamidines and matriptase or thrombin, we have docked compounds **1** and **7**, two potent inhibitors with reversed selectivity when tested for thrombin and matriptase. The goal of the docking study is to understand the structural basis of binding and selectivity of these ligands. Docking was done in two steps. First every compound was docked into the corresponding protein using the DOCK[20,21] program. During this docking the ligand was flexible while the protein structure was rigid. To take the flexibility of the protein into account, the complex structure was further optimized using MD simulation

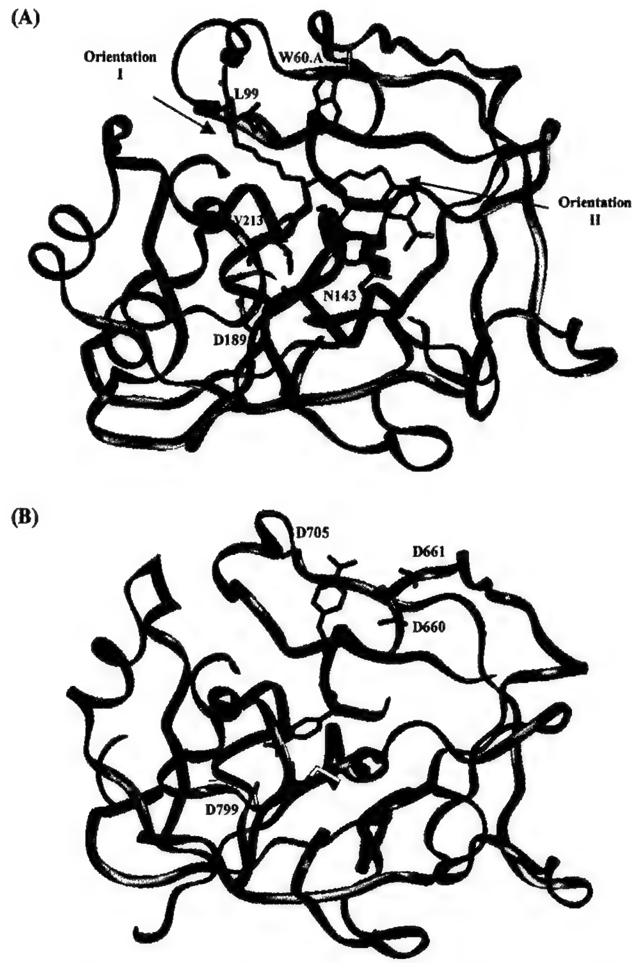


Figure 2. A: Different starting orientations of **1** in the active site of thrombin obtained after flexible ligand docking with the program DOCK. B: Orientation of **1** in the active site of matriptase obtained after flexible ligand docking with DOCK

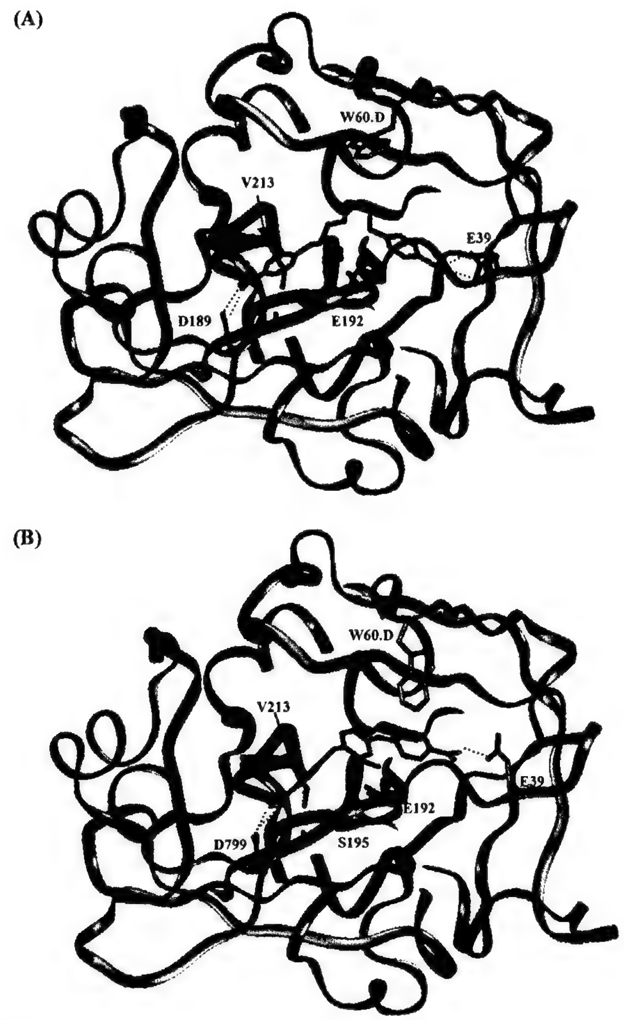


Figure 3. Lowest energy structure of the complex of compound **1** with (A) thrombin and (B) matriptase obtained after refinement using MD simulation with the generalized effective potential.

with a generalized effective potential[22] followed by conventional MD simulation. Docking with DOCK of **1** and **7** showed that these ligands adopt two different binding modes in thrombin (Figure 2A). In both binding modes, one benzamidine ring interacts with the negatively charged S1 binding site. The main difference between these two binding models is that in one binding mode (Figure 2A, orientation I) the second benzamidine fragment interacts with main-chain carbonyl oxygens in a region that corresponds to the anionic site in matriptase, while in the other binding mode (Figure 2A, orientation II), the second benzamidine fragment of the inhibitor interacts with Asn143. Since the automatic docking did not take into account the conformational flexibility of the protein, we investigated whether these two different binding models can converge to the same or similar binding model when the conformational flexibility of the protein is taken into account. For this purpose two parallel simulations were performed starting from the two different binding models of **1**. Using the MD simulation with the generalized effective potential,[22] after 1 ns, both simulations led to the same binding model of **1** in complex with thrombin, Figure 3A. Similarly, simulations of two different binding models for **7** obtained from the DOCK program led to a single converged binding model, Figure 3B. Based on the predicted binding models for **1** and **7** to thrombin, both inhibitors interact with the S1 site through a salt bridge to Asp189, hydrophobic interaction with Val213, and weak hydrogen bond between the amidino group of the inhibitor and the carbonyl oxygen of Ala190 (Figure 3A,B). While **1** interacts with Trp60.D through its linker, **7** interacts with Trp60.D through the second benzamidine fragment. While **1** forms bidentate hydrogen-bonding interactions with the carboxylate of Glu39, **7** forms only monodentate hydrogen-bonding interactions with the carboxylate of Glu39, because the rigid linker in **7** does not allow a favorable orientation of the amidino group to form a bidentate hydrogen-bonding interaction with Glu39. Bidentate hydrogen-bonding interactions between oppositely charged groups are shown to be stronger than monodentate interactions as observed for a series of thrombin inhibitors,[23] which may be one of the reasons why **7** is less potent than **1**.

Docking of **1** and **7** into matriptase using the program DOCK resulted in only one orientation for both inhibitors, Figure 2B. Both binding models were further refined using the same MD protocol as for thrombin. The refined binding models for **1** and **7** are shown in Figure 4A,B. Based upon the predicted models, **1** and **7** interact with the S1 site of matriptase through a salt bridge with Asp799, hydrogen bond with Ser800 OH, and have hydrophobic interaction with Val 824, similar to the interactions with thrombin. However, **1** forms two

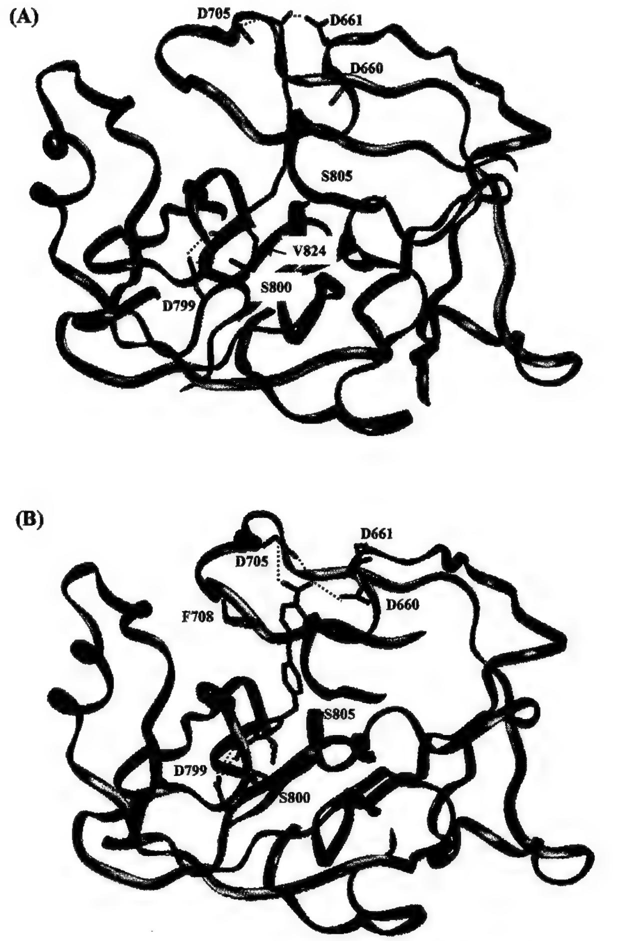


Figure 4. Lowest energy structure of the complex of compound **7** with (A) thrombin and (B) matriptase obtained after refinement using MD simulation with the generalized effective potential.

monodentate hydrogen-bonding interactions with Asp705 and Asp661 in matriptase that correspond to a bidentate hydrogen-bonding interaction with Glu39 in thrombin. However there is no hydrophobic interaction of the linker with matriptase hydrophobic residues, unlike that with thrombin. The lack of hydrophobic interactions of the linker in **1** with matriptase may be one important reason for its lower potency to matriptase as compared to thrombin. Due to its shorter linker, compound **7** was able to form a bidentate hydrogen-bonding interaction with Asp705 and a monodentate hydrogen-bonding interaction with Asp660. These interactions are stronger than that observed in thrombin, which may explain the higher potency of **7** to matriptase than to thrombin. Taken together, our docking studies provide an understanding of the structural basis of a class of novel inhibitors binding to matriptase and to thrombin and can offer a plausible explanation for the selectivity of two potent inhibitors.

Conclusion

We have identified bis-benzamidines as a class of matriptase inhibitors through structure-based database search. The lead compound, hexamidine (**1**) has K_i = 924 nM in inhibiting matriptase. Testing available analogues of the lead compound (**1**) led to the identification of **2** and **7** that are better inhibitors of matriptase with K_i = 191 and 208 nM, respectively. Compound **7** is the most selective compound for matriptase among the compounds tested. It has a selectivity of 9- and 13-fold between matriptase and uPA and between matriptase and thrombin, respectively. Our limited SAR and docking studies showed that the length of linker between the two benzamidine groups, the conformational rigidity of the linker, as well as the substituent(s) on the benzamidine ring(s) play important roles for the activity and selectivity. Differences in the relative position to the S1 site of the anionic site in matriptase versus thrombin can be used to design matriptase-selective inhibitors. The discovery of these small-molecule and nonpeptide matriptase inhibitors provides us with valuable pharmacological tools to further elucidate the biological function of matriptase. Structure-based design and chemical modifications toward improving potency and selectivity of the discovered lead compounds are currently underway and will be reported in due course.

Experimental Section

Homology Modeling and Structure Refinement. The sequence for matriptase was obtained from sequencing data. Templates for homology modeling were obtained by searching the Protein Data Bank,[14] using the program BLAST.[24] The structure of thrombin, entry 1hxe with 34% identities, 53% similarity and 6% gaps, was used as a template for modeling matriptase structure using the program MODELLER.[16] The structure obtained from homology modeling was further refined using the MD program CHARMM.[25] Hydrogen atoms were assigned to the modeled structure using the program HBUILD[26] from CHARMM. The protein was then solvated by inserting it in a 30 Å sphere of water and by deleting solvent molecules with heavy atoms that are at less than 2.5 Å from protein heavy atoms. The MD simulation was done using the all-atom parameter set from QUANTA3.2/CHARMm[27] force field, a constant dielectric $\epsilon = 1$, and constant temperature $T = 300$ K. The leapfrog method with 1-fs time step was applied for numerical integration. Long-range electrostatic forces were treated with the force switch method in a switching range of 8–12 Å. van der Waals forces were calculated with the shift method and a cutoff of 12 Å. The nonbond list was kept to 14 Å and updated heuristically. Solvent waters were kept from evaporating by using a spherical miscellaneous mean field potential as implemented in CHARMM. The solvated protein was energy minimized using 250 cycles of steepest descent and 500 cycles of adopted-basis Newton Raphson methods. This was followed by 100-ps MD simulation.

The structure of thrombin, entry 1hxe from the Protein Data Bank, was prepared for docking in the same way as the structure of matriptase.

Structure-Based Database Search. The refined structure of matriptase obtained from homology modeling, as described in the previous section, was used as the target in a structure-based 3D database search. The program DOCK[28] was used for computer-aided database screening to identify potential inhibitors. Shape and binding energy scoring functions were used to screen and rank the potential ligands. Filters were used to eliminate molecules that have more than 10 flexible bonds, to avoid considering overly flexible molecules, and also discarded molecules with fewer than 10 or more than 50 heavy atoms.

The screening of the large NCI database was done on a Silicon Graphics Indigo2 R10000 workstation. The docking was done in two stages, with the ligand flexibility being considered in both. In the first stage, two minimization cycles, with 50 iterations maximum, were considered for each compound from the database. The best-scoring 10 000 molecules were considered in the second stage, when 100 minimization cycles and 100 maximum iterations per cycle were carried out in order to refine the position of the ligand and its score. The top 2 000 compounds were then considered for selecting potential inhibitors for matriptase by inspection to determine if they contain

ionizable groups that will bind to the S1 site and the anionic site. After these screenings, 69 compounds were selected for further biochemical testing.

Molecular Modeling of the Best-Scoring Ligands. Compounds with K_i values less than 1 μM against matriptase were also docked into thrombin using DOCK with the second protocol as shown above. The orientation of the ligand obtained after docking was used as starting orientation in MD refinement, which was done on a Beowulf cluster of PCs in our laboratory. The active site of the protein–inhibitor complex was solvated by centering the molecule on Ser805/Ser195. A 20 Å radius TIP3P[29] water sphere centered in the origin was then added for solvation. Water molecules closer than 2.5 Å to any protein or inhibitor heavy atom were deleted. The solvated complex was energy minimized using 250 cycles of steepest descent and 500 cycles of adopted-basis Newton Raphson methods. This was followed by 1-ns MD refinement using the generalized effective potential implemented in CHARMM by our group.[22] The temperature of the simulation was 300 K and a 1-fs time step was used for numerical integration of the equation of motion. Long-range electrostatic forces were treated with the force switch method in a switching range of 8–12 Å. van der Waals forces were calculated with the shift method and a cutoff of 12 Å. The nonbond list was kept to 14 Å and updated heuristically. For ligand optimization an annealing protocol was used with the maximum q value[22] for calculating the generalized effective potential of 1.0005; this value was reached from the starting $q = 1$ after 10-ps simulation, followed by 10-ps simulation during which the q value was decreased to 1, which corresponds to MD, and 30-ps MD simulation. This cycle was repeated for the entire length of the simulation. A harmonic restraining force, with a force constant of 0.5 kcal/Å^2, was applied on the protein main-chain atoms that are within 20 Å of the catalytic triad Ser805/195. Residues that were farther than 20 Å from the active site Ser were fixed. This was followed by 1-ns regular MD simulation at 300 K without restraining force on the protein main-chain atoms. During this simulation residues that were farther than 20 Å from the catalytic triad Ser were also fixed. We used the same setup as for the generalized effective potential with the exception that $q = 1$ during the entire length of the simulation.

Materials. Active matriptase was purified from human milk as will be described later. Active urokinase-type plasminogen activator (uPA) was purified by aminobenzamidine-Sepharose 6B (Amersham Pharmacia, Piscataway, NJ) from a partially purified uPA from human urine. Bovin β-trypsin, bovine thrombin, and fluorescent peptide substrates *N-tert*-butoxycarbonyl (*N-t*-Boc)-Gln-Ala-Arg-7-amido-4-methylcoumarin (AMC), *N-t*-Boc-Leu-Gly-Arg-AMC, and *N-t*-Boc-Leu-Arg-Arg-AMC were purchased from Sigma (Sigma Chemical Co., St. Louis). Small-molecule inhibitors were obtained from the Drug Synthesis & Chemistry Branch, Developmental Therapeutics Program, Division of Cancer Treatment and Diagnosis, National Cancer Institute. Other chemicals are all reagent grade.

Purification of Active Form Matriptase. Activated matriptase, in a complex with its endogenous inhibitor HAI-1, was purified from human milk by immunoaffinity chromatography and maintained in its uncomplexed status in glycine buffer pH 2.4, as described previously.[11] Matriptase and HAI-1 were further separated by 10% SDS–PAGE. The proteins were stained by zinc stain kit (Bio-Rad, Hercules, CA). Gels containing the 70-kDa active matriptase were sliced out and eluted using Electro-Eluter (Bio-Rad, Hercules, CA) under nondenaturing conditions (Tris-glycine buffer, pH 8.3). Purified, active matriptase was then stored at −80 °C in acidic solution.

Determination of Inhibitory Activity. Inhibitory activity of compounds against each protease was measured at room temperature using fluorescent substrate peptides in 100 mM Tris-HCl (pH 8.5), containing 100 μg/mL bovine serum albumin. To a cuvette containing 170 μL of buffer were added 10 μL of enzyme solution and 10 μL of inhibitors. After preincubation, 10 μL of substrate was added, and the solution was mixed well by shaking the cuvette. The residual enzyme activity was then determined by following the change of fluorescence released by hydrolysis of the fluorescent substrates in a fluorescent spectrophotometer (HITACHI F4500), with excitation at 360 nm and emission at 480 nm. Peptide *N-t*-Boc-Gln-Ala-Arg-AMC was used as a substrate for matriptase and trypsin, peptide *N-t*-Boc-Leu-Gly-Arg-AMC was used as a substrate for uPA, and peptide *N-t*-Boc-Leu-Arg-Arg-AMC was used as a substrate for thrombin.

Kinetic Screening of Compound Inhibitors. The inhibitory activity of each compound was first investigated by using a fixed (75 μM) concentration both of inhibitor and matriptase. Compounds that exhibited inhibition were then subjected to a further analysis for their K_i values using Dixon plotting. We recorded the rate of hydrolysis in duplicate in the presence of 6–7 different concentrations of each inhibitor. A straight line of the concentration of inhibitor versus the reciprocal values of the rate of hydrolysis was plotted with SigmaPlot software. Two lines were obtained from two unsaturated substrate concentrations; the X value of the intersection of these lines gives the value of $-K_i$.

Acknowledgment. The chemical samples used in our biological evaluations were provided by the Drug Synthesis & Chemistry Branch, Developmental Therapeutics Program, Division of Cancer Treatment and Diagnosis, National Cancer Institute, National Institutes of Health; their help on this project is highly appreciated. Funding for this work was provided by Grants NIH 2P50CA58185 (R.B.D.) and NIH IR21-CA80897 (R.B.D. and C.-Y.L.) and by Department of Defense Fellowship DAMD 17-00-1-0269 (S.-L.L.).

References

(1) Mignatti, P.; Rifkin, D. B. Biology and biochemistry of proteinases in tumor invasion. *Physiol. Rev.* **1993**, *73*, 161–195.
(2) Westermarck, J.; Kähäri, V.-M. Regulation of matrix metalloproteinase expression in tumor invasion. *FASEB J.* **1999**, *13*, 781–792.
(3) Benaud, C.; Dickson, R. B.; Thompson, E. W. Roles of the Matrix Metalloproteinases in Mammary Gland Development and Cancer. *Breast Cancer Res. Treat.* **1998**, *50*, 97–116.
(4) Oberst, M. D.; Lin, C.-Y.; Dickson, R. B.; Johnson, M. D. Role of Proteases in Breast Cancer. *J. Women's Cancer* **2001**, in press.
(5) Takeuchi, T.; Harris, J. L.; Huang, W.; Yan, K. W.; Coughlin, S. R.; Craik, C. S. Cellular Localization of Membrane-type Serine Protease 1 and Identification of Protease-activated Receptor-2 and Single-chain Urokinase-type Plasminogen Activator as Substrates. *J. Biol. Chem.* **2000**, *275*, 26333–26342.
(6) Babine, R. E.; Bender, S. L. Molecular recognition of protein–ligand complexes: applications to drug design. *Chem. Rev.* **1997**, *97*, 1359–1472.
(7) Leung, D.; Abbenante, G.; Fairlie, D. P. Protease inhibitors: current status and future prospects. *J. Med. Chem.* **2000**, *43*, 305–341.
(8) Robinson, R. P.; Laird, E. R.; Blake, J. F.; Bordner, J.; Donahue, K. M.; Lopresti-Morrow, L. L.; Mitchell, P. G.; Reese, M. R.; Reeves, L. M.; Stam, E. J.; Yocum, S. A. Structure-based design and synthesis of a potent matrix metalloproteinase-13 inhibitor based on a pyrrolidinone scaffold. *J. Med. Chem.* **2000**, *43*, 2293–2296.
(9) Lin, C.-Y.; Wang, J.-K.; Torri, J.; Dou, L.; Sang, Q. A.; Dickson, R. B. Characterization of a novel, membrane-bound, 80-kda matrix-degrading protease from human breast cancer cells. monoclonal antibody production, isolation, and localization. *J. Biol. Chem.* **1997**, *272*, 9147–9152.
(10) Lin, C.-Y.; Anders, J.; Johnson, M.; Sang, Q. A.; Dickson, R. B. Molecular cloning of cDNA for matriptase, a matrix-degrading serine protease with trypsin-like activity. *J. Biol. Chem.* **1999**, *274*, 18231–18236.
(11) Lin, C.-Y.; Anders, J.; Johnson, M.; Dickson, R. B. Purification and characterization of a complex containing matriptase and a kunitz-type serine protease inhibitor from human milk. *J. Biol. Chem.* **1999**, *274*, 18237–18242.
(12) Lee, S.-L.; Dickson, R. B.; Lin, C.-Y. Activation of hepatocyte growth factor and urokinase/plasminogen activator by matriptase, an epithelial membrane serine protease. *J. Biol. Chem* **2000**, *275*, 36720–36725.

(13) Takeuchi, T.; Shuman, M. A.; Craik, C. S. Reverse biochemistry: use of macromolecular protease inhibitors to dissect complex biological processes and identify a membrance-type serine protease in epithelial cancer and normal tissue. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 11054–11061.
(14) Bernstein, F.; Koetzle, T. F.; Williams, G. J. B.; Meyer Jr, E. F.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Schimanouchi, T.; Tasumi, M. J. The protein data bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* **1977**, *112*, 535–542.
(15) Jonassen, I.; Eidhammer, I.; Grindhaug, S. H.; Taylor, W. R. Searching the protein structure databank with weak sequence patterns and structural constraints. *J. Mol. Biol.* **2000**, *304*, 599–619; www.rcsb.org.
(16) Sali, A.; Potterton, L.; Yuan, F.; van Vlijmen, H.; Karplus, M. Evaluation of comparative protein modeling by MODELLER. *Proteins: Struct. Funct. Genet.* **1995**, *23*, 318–326.
(17) Sali, A. Modeling mutations and homologous proteins. *Curr. Opin. Biotechnol.* **1995**, *6*, 437–451.
(18) Milne, G. W. A.; Nicklaus, M. C.; Driscoll, J. S.; Wang, S.; Zaharevitz, D. W. The NCI Drug Information System 3D Database. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1219–1224.
(19) Fersht, A. The Three-Dimensional Structure of Proteins. In *Structure and Mechanism in Protein Science. A Guide to Enzyme Catalysis and Protein Folding*; Julet, M. R., Ed.; W. H. Freeman: New York, 1999; pp 1–53.
(20) Makino, S.; Kuntz, I. D. Automated flexible ligand docking method and its application for database search. *J. Comput. Chem.* **1997**, *18*, 1812–1825.
(21) Knegtel, R. M. A.; Kuntz, I. D.; Oshiro, C. M. Molecular docking to ensembles of protein structures. *J. Mol. Biol.* **1997**, *266*, 424–440.
(22) Pak, Y.; Wang, S. Application of a molecular dynamics simulation method with a generalized effective potential to the flexible molecular docking problems. *J. Phys. Chem. B* **2000**, *104*, 354–359.
(23) Weber, P. C.; Lee, S.-L.; Lewandowski, F. A.; Schadt, M. C.; Chang, C.-H.; Kettner, C. A. Kinetic and crystallographic studies of thrombin with Ac-(D)Phe-Pro-boroArg-OH and its lysine, amidine, homolysine, and ornithine analogues. *Biochemistry* **1995**, *34*, 3750–3757.
(24) Altschul, S. F.; Madden, T. L.; Schäffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. Gapped BLAST and PSI–BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402.
(25) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **1983**, *4*, 187–217.
(26) Brunger, A. T.; Karplus, M. Polar hydrogen positions in proteins: empirical energy placement and neutron diffraction comparison. *Proteins: Struct. Funct. Genet.* **1988**, *4*, 148–156.
(27) Momany, F. A.; Rone, R. Validation of the general purpose QUANTA 3.2/CHARMm force field. *J. Comput. Chem.* **1992**, *13*, 888–900.
(28) Bemis, G. W.; Kuntz, I. D. A fast and efficient method for 2D and 3D molecular shape description. *J. Comput.-Aid. Mol. Des.* **1992**, *6*, 607–628.
(29) Durell, S. R.; Brooks, B. R.; Ben-Naim, A. Solvent-induced forces between two hydrophilic groups. *J. Phys. Chem.* **1994**, *98*, 2198–2202.

JM000395X

Submitted to: Bioorg & Medicinal Chem. Lett. on Jan 31, 2001

Synthesis and Evaluation of the Sunflower Derived Trypsin Inhibitor as a Potent Inhibitor of the Type II Transmembrane Serine Protease, Matriptase

Ya-Qiu Long,[a] Sheau-Ling Lee,[b,c] Chen-Yong Lin,[b] Istvan Enyedy,[b] Shaomeng Wang,[b] Robert B. Dickson,[b] and Peter P. Roller,[a,*]

[a]Laboratory of Medicinal Chemistry, National Cancer Institute, NIH, FCRDC, 376/208, P.O. Box B, Frederick, MD 21702, USA; [b]Lombardi Cancer Center, Georgetown University Medical Center, Washington, DC 20007, USA

Abstract: We report here the synthesis of a 14 amino acid long bicyclic peptide, previously isolated from sunflower seeds. This peptide, termed sunflower trypsin inhibitor (SFTI-1), is one of the most potent naturally occurring small-molecule trypsin inhibitors. It is comprised of a backbone cyclized loop bisected with a disulfide bridge. In addition to inhibiting trypsin, the synthetic SFTI-1 is also a very potent inhibitor, with a K_i of 0.92 nM, of the recently derived epithelial serine protease, termed 'matriptase'.

Introduction: Plant derived protease inhibitors serve in the defense mechanisms of plants against pests and plant pathogens (1). These inhibitors can be classified into a number of families based on their active-site structures and their specificities to inhibit the cleavage of specific peptide sequences within proteins. The majority of these inhibitors are classified as serine proteases. One of the well known serine protease inhibitory agents, the Bowman-Birk inhibitor, found in seeds of legumes and other plants, belongs to the Bowman-Birk inhibitor family of small proteins with a MW range of 6000-8000 (2). These proteins inhibit trypsin, chymotrypsin or elastase, depending on the configuration of the reactive site loop within the inhibitor.

Recently, a 14-amino acid peptide, termed sunflower trypsin inhibitor (SFTI-1) was isolated from sunflower seeds (3). SFTI-1 inhibited β-trypsin with an impressive subnanomolar K_i of 0.1 nM, and it inhibited cathepsin G with a comparable K_i (3). SFTI-1 has considerable selectivity; for example, it proved to be 74-fold less inhibitory for chymotrypsin, and was found to be 3 orders of magnitude less inhibitory for elastase and thrombin. In contrast, it had no effect on Factor Xa (3).

---

*Corresponding Author: Tel: +1-301-846-5904; Fax: + 1-301-846-6033; email: proll@helix.nih.gov

[c]Supported by Department of Defense Fellowship DAMD 17-00-1-0269

The natural product, SFTI-1 was partially characterized by classical techniques, and its structure confirmed, based on the electron density map of the inhibitor co-crystallized with bovine-β-trypsin (3). SFTI-1 is one of the smallest, naturally occurring plant protein inhibitors reported to date; it has considerably enhanced potency, relative to other peptides of similar length. Its backbone-cyclized peptide structure is additionally stabilized by a cystine disulfide bond. We report here the facile synthesis of SFTI-1. This methodology will also provide for the synthesis of various analogs, with altered inhibitory profiles, relative to various serine proteases of interest.

One such serine protease of interest, is known as matriptase (or MT-SP1), a member of the emerging class of type II transmembrane serine proteases (4,5,6). The mouse homolog of matriptase has also been described and is termed epithin (7). Matriptase /epithin is of considerable interest for the development and pathogenesis of epithelial tissues. Although matriptase is initially synthesized by multiple types of epithelial cells as a transmembrane serine protease, it was isolated originally from human milk in its activated form complexed with its cognate Kunitz type of serine inhibitor (KSPI), the hepatocyte growth factor activator inhibitor (HAI-1). Additionally, matriptase was isolated from human breast cancer cells in culture (4,5,8). Matriptase may function to degrade the extracellular matrix, as well as several cellular regulatory proteins; specifically, it may activate hepatocyte growth factor (HGF) by cleaving its inactive proform, it may activate urokinase by cleaving its zymogen, and it may cleave and activate the protease activated receptor-2, PAR-2 (9,10). Thus, matriptase blockade could potentially modulate cell proliferation, motility, invasion, and differentiation of cells (6,10). We report here the synthesis of SFTI-1 and its potent enzyme inhibitory effectiveness with matriptase, in comparison to its potency for related serine proteases. SFTI-1 thus represents a potentially useful inhibitor for biological studies and therapy of diseases, including cancer.

Synthesis of SFTI-1

Fmoc chemistry based solid phase peptide synthesis methodology was used for the synthesis of the linear peptide RCTKSIPPICFPDG-Rink resin. The total synthesis of bicyclic peptide SFTI-1 is described in Scheme 1. The acid sensitive 4-(2',4'-dimethoxyphenyl-

hydroxyphenyl)-Phenoxy('Rink resin') was purchased from Bachem California Inc. (Torrance, CA, USA). Fmoc derivatives of standard amino acids were obtained from Perkin-Elmer/Applied Biosystems Division. Side-chain protections are as follows: Arg(Pmc), Cys(Trt), Thr(t-Bu), Lys(Boc), Ser(t-Bu), Asp(OtBu). HBTU/HOBt activation of N^{α}-protected amino acids was employed for coupling, and 20% piperidine/DMF was used for Fmoc deprotection. HATU/HOAt/DIEA in DMF was used for backbone cyclization. The crude peptide was purified by reverse-phase high-performance liquid chromatography (RP-HPLC). HPLC conditions: Vydac C4 column (20 X 250 mm); solvent gradient, A, 0.05% TFA in water; B, 0.05% TFA in 90% acetonitrile in water with gradient indicated below; flow rate, 10 mL/min; UV detector, 225 nm. FAB-MS (unit resolution, glycerol matrix) was performed on a VG Analytical 7070E-HF mass spectrometer. Amino acid analysis (6N HCl, 110°C, 24 h) was carried out at the Protein and Carbohydrate Structure Facility (University of Michigan, Ann Arbor, Michigan, USA).

The first amino acid, glycine, was attached to the acid labile Rink resin. Positioning Gly at the C-terminal end minimizes the possibility of racemization on subsequent backbone cyclization. Fmoc-Gly-OH (297 mg, 1.0 mmol) was double coupled with the Rink resin (357 mg, 0.1 mmol) by using 1,3-diisopropylcarbodiimide (1 mmol), 4-(dimethylamino)pyridine (0.1 mmol) and N-methylmorpholine (0.1 mmol) in 3 mL DMF (2 hr, RT). The Fmoc-Gly-Rink resin was dried under a vacuum and treated as a preloaded resin for automated peptide synthesis with an ABI 433A peptide synthesizer using the FastMoc protocol. On completion of the sequence RCTKSIPPICFPDG-Rink resin, the side-chain protected peptide was cleaved from the resin with 20 mL of 1% TFA/DCM (1 min at RT). The solution was collected by filtration and neutralized with NMM in an ice bath. The resin was treated similarly 2 more times with 1% TFA/DCM (5 min at RT) and the combined neutralized solution was evaporated to dryness *in vacuo*. For backbone cyclization, the crude side-chain protected peptide was dissolved in anhydrous DMF (80 mL), and added DIEA (0.6 mmol) and HATU/HOAt (0.30 mmol each) at ice bath temperature. The reaction mixture was stirred at RT for 24 h, neutralized with 30% AcOH and evaporated *in vacuo*. The residue was treated with 95% TFA containing 2.5% each of triethylsilane (TES) and H_2O to deprotect the side chains. Subsequently, for disulfide oxidization, the head-to-tail cyclic peptide was dissolved in 20 mL of water and added dropwise to 250 mL of water solution, which was previously adjusted to pH 8.5 with ammonium acetate

and ammonium hydroxide. The disulfide cyclization occurred spontaneously under the basic conditions while bubbling with oxygen for 6 h, and then stirring overnight at RT. The reaction was quenched by adding AcOH and the solution was lyophilized. The synthetic SFTI-1 was purified by RP-HPLC, Rt = 17.6 min (gradient 20-60% B over 40 min); FAB-MS $(M+H)^+$ 1513.0 (calc. 1513.7). Amino acid analysis: Asp 1.02(1), Ser 0.91(1), Thr 1.15(1), Pro 2.79(3), Lys 0.97(1), Gly 1.15(1), Arg 1.16(1), Ile 1.83(2), Phe 1.02(1).

[Scheme 1]

Enzyme Inhibitory Assays

The 70-kDa activated matriptase was isolated as described previously (5,7,8). Urokinase-type plasminogen activator (uPA) was purified by aminobenzamidine-Sepharose 6B (Amersham Pharmacia, Piscataway, NJ) from a partially purified uPA from human urine. Bovine β-trypsin, bovine thrombin, Bowman-Birk Inhibitor (BBI), and the fluorescent substrates were purchased from Sigma (Sigma Chemical Co., St. Louis). Inhibitory activity of SFTI-1 to proteases was measured at RT in a reaction buffer of 100 mM Tris-HCl (pH 8.5) containing 100 mg/mL of bovine serum albumin, using the fluorescent substrate peptides. To a cuvette containing 170 μL of reaction buffer was added 10 μL of enzyme solution and 10 μL of inhibitor solution. After preincubation, 10 μL of substrate solution was added and the cuvette content mixed thoroughly. The residual enzyme activity was determined by following the change of fluorescence released by hydrolysis of the substrates, using a fluorescent spectrophotometer (HITACHI F4500) with excitation wavelength of 360 nm and emission at 480 nm. Fluorescent peptide N-t-Boc Gln-Ala-Arg-AMC was used as substrate for matriptase and trypsin, peptide N-t-Boc-Leu-Gly-Arg-AMC was used as substrate for uPA, and peptide N-t-Boc-Leu-Arg-Arg-AMC was used as substrate for thrombin. Hydrolysis rates were recorded in presence of 6-7 different concentrations of SFTI-1. The K_i values were determined by Dixon plots from 2 sets of data with different concentrations of substrate.

[Table 1]

Modeling of SFTI-1 in Complex with Matriptase

The homology modeling approach was used to build the 3D structure of matriptase. A search of the protein databank (12), using the program BLAST (13), has identified thrombin (PDB entry: 1hxe) as a good template for homology modeling, with 34% identities, 53% positives, and 6% gaps. The structure of matriptase was built using the program MODELLER (14). The structure derived in this manner was refined by its solvation with TIP3P water molecules, minimization to eliminate bad contacts, and equilibration at 300 K using molecular dynamics simulation with the program CHARMM (15). For docking purposes, only the active site of matriptase was solvated with a 20 Å radius water sphere of TIP3P water, with the center defined by the average of SFTI-1 Cys3 S, Ser191 O, Ser191 C, and Ser191 C. Only the surface residues of matriptase were considered flexible, and the rest were fixed during the simulation. The structure of SFTI-1 was obtained from the protein databank (PDB entry: 1sfi). The starting orientation of the inhibitor is the same as the one found in the X-ray structure with trypsin, 1SFI (3). The docking of matriptase with the inhibitor was performed using molecular dynamics with Tsallis effective potential (16,17), as implemented in CHARMM (15). The simulation was done using the all atom parameter set CHARMM22, the temperature was set to 300 K, the time step for integration was 1 fs, the potential shift parameter was 11000 kcal/mol, and q, the final value of the Tsallis coefficient, was 1.001. The q-annealing schedule was set to 20 ps, for increasing q from 1.0 to 1.001, 20 ps for decreasing q from 1.001 to 1.0, and 10 ps normal molecular dynamics. The total length of simulation was 4 ns. Long-range electrostatic forces were treated with the force switch method in a switching range of 8-12 Å. Van der Waals forces were calculated with the shift method and a cutoff of 12 Å. The nonbond list was kept to 14 Å, and updated heuristically. Solvent waters were kept from evaporating by using a spherical miscellaneous mean field potential as implemented in CHARMM. An NOE restraint was used on the distance between SFTI-1 Lys5 N and Asp185 C of the protein. This restraint kept Lys5 of SFTI-1 bound to the S1 pocket of matriptase; Asp185 is at the bottom of the binding pocket. The maximum value of the distance was set to 4 Å, similar to the distance in the X-ray structure of trypsin complexed with SFTI-1 (3).

Discussion

Endogenous proteases play a pivotal role in the normal cellular physiology of the cell, such as the proteolytic activation of peptide hormones and the activation of message-transmitting peptides and proteins (18). Our efforts are focussed on regulation of a recently identified protease, termed matriptase. Matriptase was initially purified from human milk, but it is also produced by normal and cancerous epithelial cells in culture. Recent studies have suggested that inappropriate expression of its active form has the potential of producing deleterious effect in tissues, contributing to pathogenic states, such as cancer (5,8). This particular serine protease can degrade extracellular matrix proteins, and activate specific proteins, such as HGF, uPA, and PAR-2, by cleaving its inactive pro-form (9,10). These effects are likely to contribute to abnormal cell proliferation, motility, and states of differentiation. In efforts to identify inhibitors of matriptase for experimental purposes, we evaluated the recently identified sunflower derived trypsin inhibitor (SFTI-1), and found it to be a highly effective inhibitor of the enzyme.

SFTI-1 was recently isolated from sunflower seeds, as a complex with trypsin, and its structure was determined by NMR spectroscopy, and by X-ray crystallography (3). Its promising protease inhibitory profile, reported in the original work (3), prompted us to develop the current synthetic methodology for preparation of this 14-amino acid bicyclic peptide, as outlined in Scheme 1. The amino acid backbone was assembled on Rink resin. After resin cleavage, the fully sidechain-protected peptide was backbone cyclized, followed by deprotection of all sidechains. Air oxidation, in weakly basic medium, smoothly provided in good yield the intramolecularly bridged cyclic peptide, SFTI-1. NMR and X-ray studies demonstrated that SFTI-1 has considerable structural rigidity (3), imparted by the intramolecular disulfide bond. The overall conformation and the amino acid sequence was very similar to the reported structures for the equivalent trypsin-inhibitory loop of the reactive site segments within the Bowman-Birk inhibitors (BBI). That segment (-CTKSIPP-) is also very similar to the trypsin inhibitory segment of the mung bean inhibitory protein (19).

[Figure 1.]

Table 1 summarizes the results for a relevant set of serine proteases. The synthesized SFTI-1, just as its natural form, is a potent inhibitor of trypsin with a Ki value of 1.06 nM. In addition, we found that SFTI-1 inhibited matriptase with comparable effectiveness to trypsin. Although the modeled 3D structure of matriptase was built using thrombin as template, SFTI-1 is a poor inhibitor of thrombin (K_i = 5,050 nM). SFTI-1 is also non-inhibitory for uPA, an important serine protease in the extracellular matrix degradative network. This selectivity of SFTI-1, thus, makes it a valuable tool to study the function of matriptase in biological matrices.

On account of the overall similarity of the secondary structure and 3D structure between SFTI-1 and the trypsin inhibitory loop of BBI, we tested the inhibitory activity of BBI to matriptase. A shown in Figure 1, at concentrations of 1.25, 2.5 and 5 μM, BBI significantly reduced the matriptase hydrolytic activity of standard substrates, compared with the SFTI-1 at the same concentration; however, the inhibitory potency of SFTI-1 is 2-3 fold higher than that of BBI.

[Figure 2, Molecular Model]

A molecular modeling study was initiated in order to better understand the structural features that contribute to the high inhibitory activity of SFTI-1. The predicted structure of matriptase complexed with SFTI-1 was obtained after docking using 4 ns molecular dynamics simulations with the Tsallis effective potential, as shown in Figure 2. This structure was compared to the crystal structure of trypsin, in complex with SFTI-1 (3), to investigate the similarities and differences between the interactions of SFTI-1 with these two proteins. In both structures, Lys5 of the SFTI-1 binds to the S1 site of protease; interacting with Asp189 in trypsin or Asp799 in matriptase. Arg2 of the SFTI-1 mainly interacts with the backbone carbonyl group of Asn97 and solvent molecules in the X-ray structure. Arg2 forms an H-bond with the Phe706 main chain carbonyl and interacts with Phe708 and Phe706 sidechains of matriptase through π - cation interactions. Asp14 in SFTI-1 is largely exposed to solvent in both structures; it forms a hydrogen bond with Asn72 in the trypsin X-ray complex structure, but mainly interacts with Arg2 of SFTI-1 in the docked matriptase structure. Phe12 of SFTI-1 is in close proximity with only one hydrophobic residue, Leu99, in the trypsin complex structure, while this residue interacts with Phe708 in the docked complex structure with matriptase. Our docking studies

suggest that SFTI-1 has similar interactions with matriptase and trypsin, and these results provide a rational explanation for the comparative inhibitory potency of SFTI-1 to these two proteases.

The constrained, conformationally rigid structure of the bicyclic peptide, SFTI-1, provides for a promising pharmacophore model towards further development of more specific inhibitors of extracellular matrix serine-proteases, such as matriptase. Synthetic methodologies allow for the design and synthesis of various analogs, based on the homology-modeled matriptase catalytic site. Synthetic approaches are also available now for the design and generation of libraries of back-bone cyclized peptides with bisecting cystine disulfide bridges, as demonstrated for the development of enzyme inhibitory BPTI mimetics (20).

References and Notes

1. Shewry, P. R.; Lukas, J. A. *Advan. Botan. Res.* **1997**, *26*, 135-192.
2. Laskowski, M.; Kato, L. *Annu. Rev. Biochem.* **1980**, *49*, 593-626
3. Luckett, S.; Santiago Garcia, R.; Barker, J. J.; Konarev, A. V.; Shewry, P. R.; Clarke, A. R.; Brady, R. L. *J. Mol. Biol.* **1999**, *290*, 525-533.
4. Lin, C.-Y.; Wang, J.-K.; Torri, J.; Dou, L.; Sang, Q. X. A.; Dickson, R. B. *J. Biol. Chem.* **1997**, *272*, 9147-9152.
5. Lin, C.-Y.; Anders, J.; Johnson, M.; Sang, Q. A.; Dickson, R. B. *J. Biol. Chem.* **1999**, *274*, 18231-18236.
6. Hooper, J. D.; Clements, J. A.; Quigley, J. P.; Antalis, T. M. *J. Biol. Chem.* In Press.**(2001)**.
7. Kim, M. G.; Chen, C.; Lyu, M. S.; Cho, E. G.; Park, D.; and Schwartz, R. H. *Immunogenetics* **1999**, *49,* 420-428.
8. Lin, C.-Y.; Anders, J.; Johnson, M.; Dickson, R.B. *J. Biol. Chem.* **1999**, *274*, 18237-18242.
9. Takeuchi, T., Harris S.; Hwang, W.; Yan, K. W.; Coughlin, S. R.; and Craik, C. S. *J. Biol. Chem.* **2000**, *275*, 26333-26342.
10. Lee, S.-L.; Dickson, R. B.; Lin, C. Y. *J. Biol. Chem.* **2000**, *275*, 36720-36725.
11. Kawaguchi, T.; Qin, L.; Shimomura, T,; Kondo, J.; Matsumoto, K.; Denda, K.; Kitamura, N. *J. Biol. Chem.* **1997**, *272*, 27558-27564.
12. (a) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B., Mayer, Jr., E. F.; Brice, J. M. D.; Rogers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. *J. Mol. Biol.* **1977**, 112, 535-542. (b) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235-242.
13. Altschul, S. F.; Madden, T. L.; Schäffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. *Nucleic Acids Res*. **1997**, *25*, 3389-3402.
14. Sali, A.; Potterton, L.; Yuan, F.; van Vlijmen, H.; Karplus, M. *PROTEINS: Structure, Function, and Genetics* **1995**, *23*, 318-326.
15. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187-217.
16. Andricioaei, I.; Straub, J. E. *Physical Rev. E.* **1996**, *53*, R3055-R3058.
17. Pak, Y.; Wang, S. *J. Phys. Chem. B.* **2000**, *104*, 354-359.

18. Wenzel, H. R.; Tschesche, H. In *Peptides: Synthesis, Structure and Applications;* Gutte, B., Ed.; Academic Press: San Diego, CA, 1995; pp 321-362.

19. Li, Y.; Huang, Q.; Zhang, S.; Liu, S.; Chi, C.; Tang, Y. *J. Biochem.* **1994**, *116*, 18-25.

20. Kasher, R.; Oren, D. A.; Barda, Y.; Gilon, C. *J. Molec. Biol.* **1999**, *292*, 421-429.

Table 1. Protease Inhibitory properties of SFTI-1

Protease	K_i (nM)[a]
Matriptase	0.92
Trypsin	1.06[b]
Thrombin	5,050
uPA	500,000

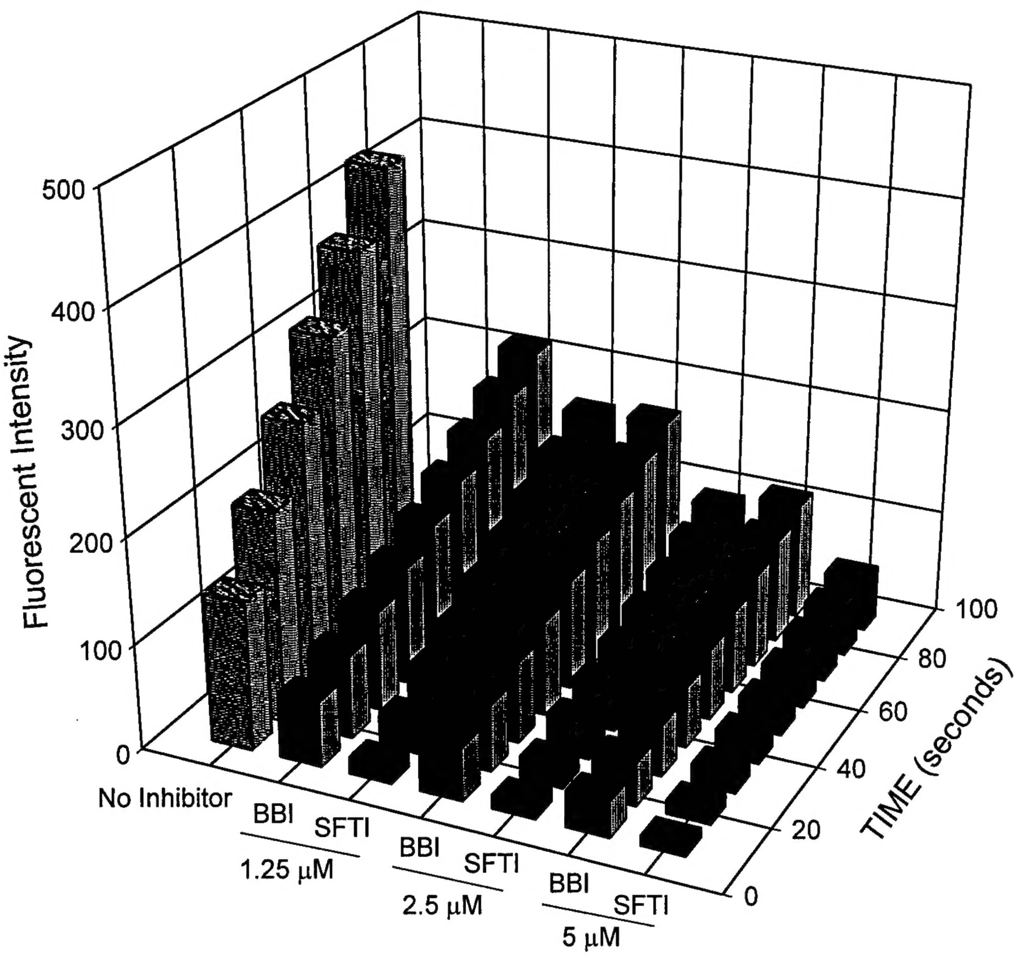
[a]Assay conditions as described in 'Enzyme Inhibitory Assays' section.

[b]Note: Ki of 0.1 nM was reported by Luckett et al (3) using competitive assay conditions.

Figure Legends

Figure 1. SFTI-1 is several fold more potent inhibitor of matriptase than BBI at low μ-molar concentrations. Figure shows the released fluorescence of the peptide substrate resulting from the proteolytic cleavage by matriptase. The assays were carried out as described in the 'Enzyme Inhibitory Assays' section, in the absence of inhibitors (hatched bars), or in the presence of various concentrations of SFTI-1 (black bars) or of BBI (gray bars).

Figure 2. Modeling based structure of matriptase complexed with SFTI-1. For details, see the Experimental Section.



Fluorescent Intensity
500
400
300
200
100
0
No Inhibitor
BBI
SFTI
1.25 μM
BBI
SFTI
2.5 μM
BBI
SFTI
5 μM
TIME (seconds)
0
20
40
60
80
100

Fig. 1

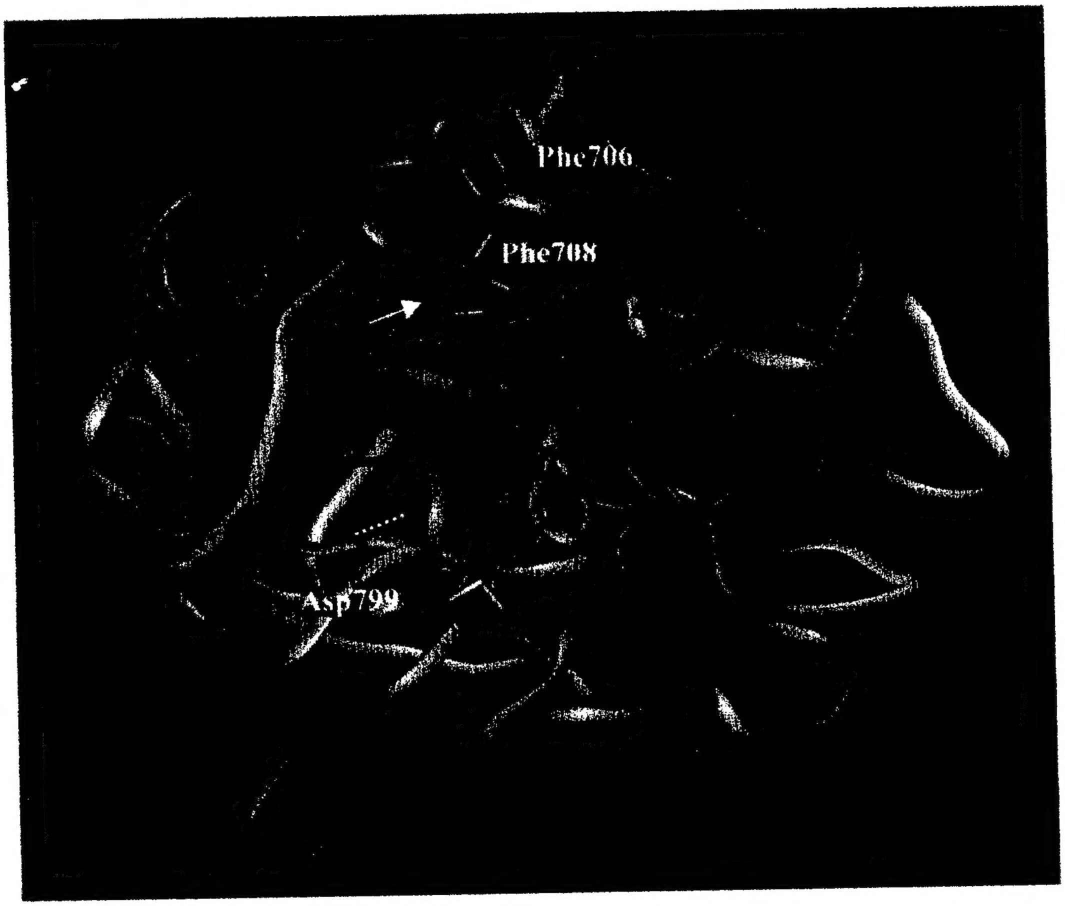
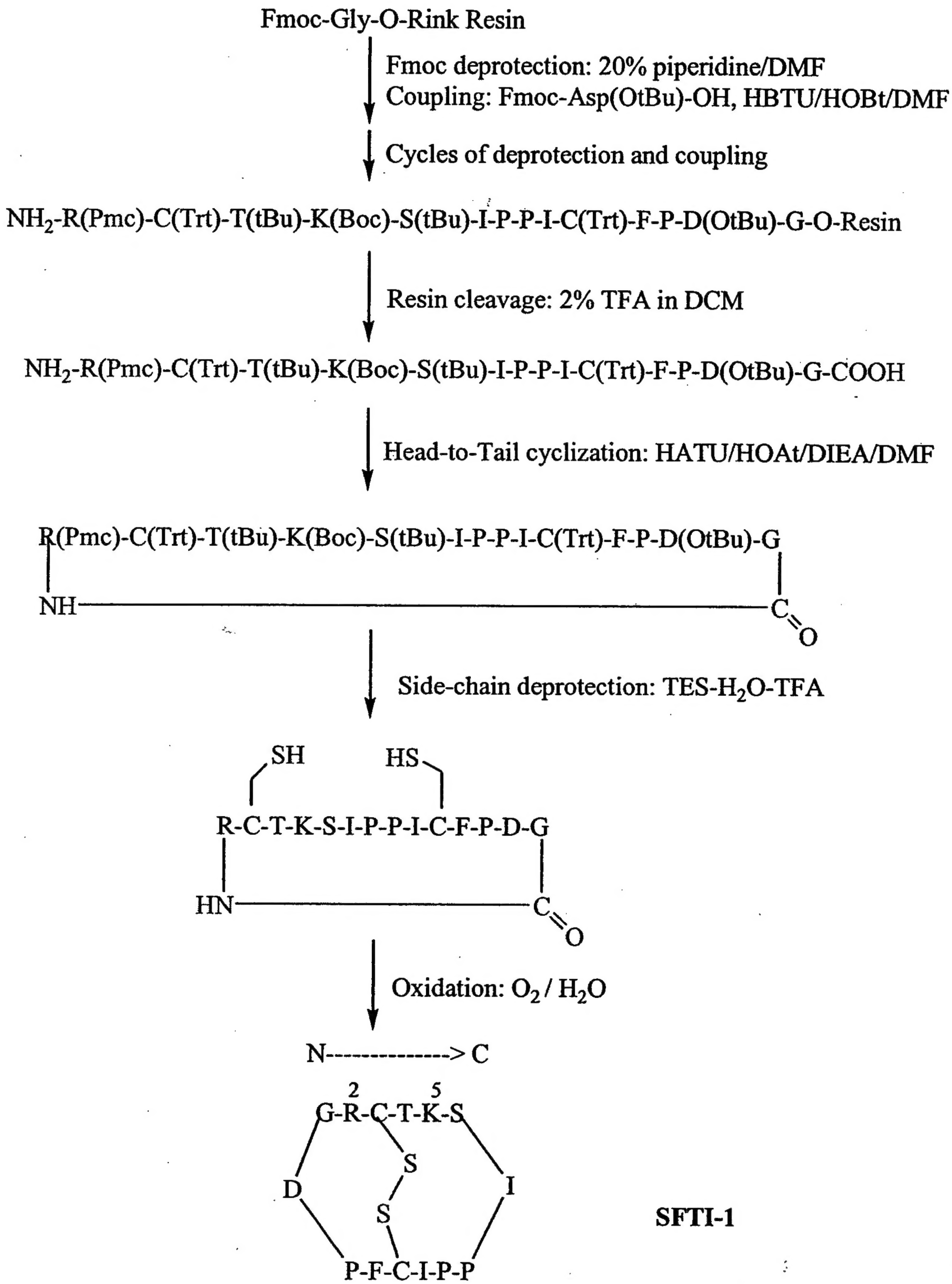


Fig 2


Fmoc-Gly-O-Rink Resin
Fmoc deprotection: 20% piperidine/DMF
Coupling: Fmoc-Asp(OtBu)-OH, HBTU/HOBt/DMF
Cycles of deprotection and coupling
NH2-R(Pmc)-C(Trt)-T(tBu)-K(Boc)-S(tBu)-I-P-P-I-C(Trt)-F-P-D(OtBu)-G-O-Resin
Resin cleavage: 2% TFA in DCM
NH2-R(Pmc)-C(Trt)-T(tBu)-K(Boc)-S(tBu)-I-P-P-I-C(Trt)-F-P-D(OtBu)-G-COOH
Head-to-Tail cyclization: HATU/HOAt/DIEA/DMF
R(Pmc)-C(Trt)-T(tBu)-K(Boc)-S(tBu)-I-P-P-I-C(Trt)-F-P-D(OtBu)-G
NH
C
O
Side-chain deprotection: TES-H2O-TFA
SH
HS
R-C-T-K-S-I-P-P-I-C-F-P-D-G
HN
C
O
Oxidation: O2 / H2O
N--------------> C
2
5
G-R-C-T-K-S
S
D
I
S
P-F-C-I-P-P
SFTI-1

Scheme 1

Scheme 1. Synthetic route for sunflower trypsin inhibitor **SFTI-1**

Scheme (Legend)